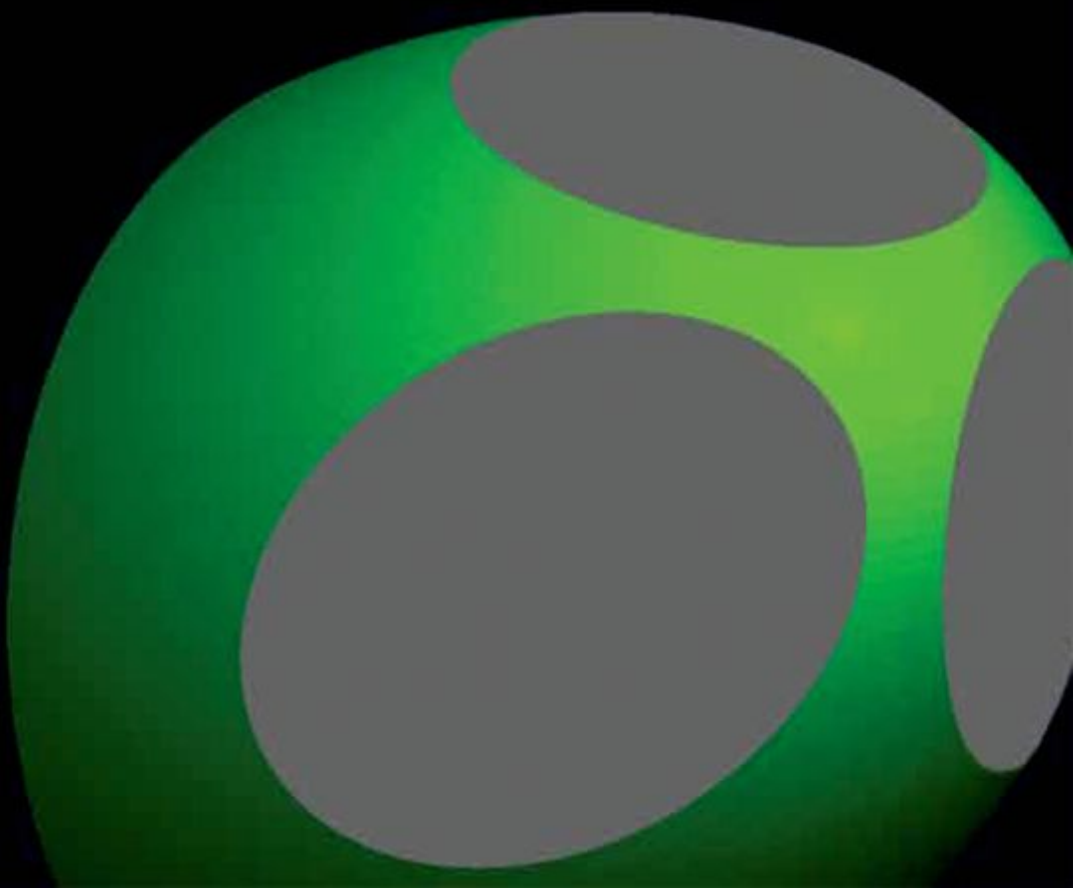
OXFORD

The Theory of Materials Failure

RICHARD M. CHRISTENSEN



The Theory of Materials Failure

This page intentionally left blank

The Theory of Materials Failure

Richard M. Christensen
Stanford University

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Great Clarendon Street, Oxford, OX2 6DP,
United Kingdom

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide. Oxford is a registered trade mark of
Oxford University Press in the UK and in certain other countries

First Edition published in 2013

Impression: 1

British Library Cataloguing in Publication Data
Data available

ISBN 978–0–19–966211–1

Printed in Great Britain by
CPI Group (UK) Ltd, Croydon, CR0 4YY

Preface

The technical and scholarly interest in materials failure goes back almost to the beginnings of classical mechanics and deformable body mechanics. The effort to put order and organization into the field of failure characterization and failure criteria has been unflagging over the ensuing time span, measured in decades and even centuries. Despite the high level of sustained activity, the long time rate of progress was agonizingly slow.

By many measures of difficulty, the treatment of failure in solids (materials) is comparable to that of turbulence in fluids, both being controlled by non-linear physical effects. It is only in the modern era that the elements needed for constructing a complete, three-dimensional theory of failure for homogeneous materials have coalesced into meaningful forms. This book presents the derivation and a detailed examination of the resultant general theory of failure for materials science and materials engineering.

Chapter 1 outlines the materials failure problem and completely sets the course for all that follows. The coverage spans the full range, starting with the efforts from some of history's greatest scientists and ultimately leading to the most recent developments, such as the failure of anisotropic fiber composite materials and an examination of microscale and nanoscale failure. Many interrelated areas of the materials failure discipline are included. Although the coverage is broad, there is no compromise in quality or rigor.

The world of materials synthesis and materials applications offers many opportunities and many challenges. Few are of higher priority or in greater need than that of understanding materials failure.

This page intentionally left blank

Contents

Recognition

This book could not have been written without the supporting foundations of long-established elasticity theory and modern plasticity theory. In this connection, a special note of remembrance and appreciation is due to

Rodney Hill
and
Daniel C. Drucker

Their individual efforts (along with others) gave substance and clarity to the plasticity formalism, and greatly helped to solidify it as a major discipline alongside that of elasticity. With elasticity theory and plasticity theory firmly in place, failure theory could complete the trilogy. Perhaps this would have pleased Professor Drucker and Professor Hill, since each of them took initial steps in that direction.

Technical Status and Challenges

A complete and comprehensive theory of failure for homogeneous materials is developed. The resulting general failure theory for isotropic materials and its related failure criteria are calibrated by two properties: the uniaxial tensile and compressive strengths, T and C. From such readily available data for most materials, the entire family of failure envelopes can be generated for any and all states of stress in any isotropic material. It is not just a coincidence that the number of independent failure properties being at two is the same as the number of independent elastic properties for isotropy. This relationship will be of considerable consequence.

It will require a long and involved derivation to establish the operational capability and results mentioned briefly above. First, however, a historical survey and evaluation must be conducted. Insofar as general applications are concerned, the complete and absolute unsuitabilities are detailed for the Mises, Tresca, Drucker–Prager, and Coulomb–Mohr failure criteria. Only the Mises form is satisfactory for a specific class of materials—that being only for ductile metals. A general failure theory must cover not only ductile metals but also brittle metals, glassy and crystalline polymers, ceramics, glasses, and a variety of other isotropic materials types, just as does elasticity theory. Treating failure criteria with the appropriate and necessary generality has always remained an unsolved and historically formidable problem, and significant progress has been greatly impeded—even blocked.

The failure problem and its need for resolution provides the impetus for the account and developments presented here. The end result is the long-missing general theory of materials failure, contained herein. All of this is made possible by a new and transparently clear physical insight or recognition. It is that of an organizing principle whereby the entire spectrum of isotropic materials can be characterized by and classified by their uniaxial strengths ratio, T/C. The limits on T/C define the ductile and brittle limits of physical behavior. Two coordinated but competitive failure criteria can then be derived: the polynomial-invariants criterion and the fracture criterion—both expressed in terms of T/C and

with stress non-dimensionalized by C. This formalism identifies universal failure forms applicable to all materials at the same T/C specification. It is a complete and self-contained theory, secured by the two most basic strength properties conceivable. It is the field theory of failure.

One of several unique features of the book is a thorough treatment of ductile versus brittle behavior for isotropic materials. The fundamental form for the ductile/brittle transition is derived as part of the failure theory. A means for gauging ductility levels is a further outcome from the derivation. Along with the experimental evaluation and verification of the failure theory, many examples of failure behavior and applications of failure criteria are presented.

The relationship of failure criteria (for homogeneous materials) to fracture mechanics (for structures) is included. Rigorous definitions of yield stress and strength are developed. Reasonably general treatments are presented for anisotropic fiber composites failure as well as investigations into microscale and nanoscale aspects of failure. Last, but still very important, there is a probe into damage models leading to failure and a fairly extensive derivation of probabilistic failure and probabilistic life prediction to complete the book. All of these technical areas are difficult and important failure-related problems in their own right.

The intended use for the book is as follows. Although it is at a quite high level, it fits in with mainstream mechanics of materials curricula in materials science and in mechanical, aerospace, and civil engineering departments. It is thus seen to be appropriate for both classroom use (upper undergraduate or graduate) and for related research. Despite the advanced level, the final forms of the failure criteria are totally practical and adaptable for use by engineers having design responsibilities.

I am deeply appreciative of Stanford University and the Office of Naval Research for creating and sustaining an environment that encourages the best possible work. To a lifetime of friends and colleagues, I am grateful for the opportunity to have interacted with them in such important, exciting, and exacting fields.

Richard M. Christensen
Aeronautics and Astronautics Department
Stanford University

This page intentionally left blank

1
The Perspective on Failure and Direction of Approach

This book is concerned with the means and methods of three-dimensional analysis for predicting materials failure. It is the first book on theoretically derived but physically based approaches for predicting the envelopes of failure for all the major classes of materials. Such problems have almost always been treated by empirical means. The approach uses the intrinsic static strength properties for homogeneous materials to predict their failure behavior in the complex multi-dimensional stress conditions of physical applications.

1.1 Materials Failure Problem

The use of failure criteria has the same motivation as that when fracture mechanics is used to predict the failure of structures due to extreme stress concentration conditions at flaws in their load-bearing elements. In fact, the failure criteria of concern here for homogeneous materials and fracture mechanics comprise complementary disciplines. Each is intended to treat a distinct and extremely broad class of problems.

The title of this book, *The Theory of Materials Failure*, has a meaningful relationship with the titles and contents of the longstanding books [1.1]–[1.3]:

- *Theory of Elasticity*, Timoshenko and Goodier
- *Mathematical Theory of Elasticity*, Sokolnikoff
- *The Mathematical Theory of Plasticity*, Hill

This book presumes familiarity with the notation and general coverage of the above books, or other more recent but largely equivalent books. More than that, it is the logical continuation and conclusion of these and other similar books, in the following sense. Elasticity and plasticity

provide part of the constitutive formulation for materials behavior. The other part is the missing part; namely, what is the permissible range for these behaviors? The range of behavior is restricted by failure, and failure criteria are the related governing constitutive forms. In briefest, broadest terms, this book is intended to supply the missing physical and mathematical complement to all elasticity and plasticity books, especially including the three classics noted above.

Specific to the coordinated opportunity represented by failure theory is the example of elasticity theory itself. The tensor-based theory of elasticity has had a revolutionary effect on the way materials are made into load-bearing structures. Elasticity is one of the oldest and most valued of all field theories, and is applied as a standardized and required procedure in virtually all applications. A tensor-based theory of failure would be the full partner to elasticity theory. Failure characterization can be seen as the three-dimensional completion and terminus of linear elastic behavior. Plasticity theory and behavior does not contradict this; rather, it represents a more complex transition from the elastic state to failure in the case of ductile materials. All of this is symptomatic of the physical problem itself: the failure of materials due to excessive states of loading.

With or without plasticity, failure still essentially represents the terminus of the elastic behavior. This will turn out to be a pivotal key to the problem. In general, elasticity, plasticity, and failure theories together comprise the continuum of materials responses to states of imposed stress. Each requires its own constitutive formulation. Two of the three corresponding knowledge bases are complete and standardized. The third—failure—is strongly needed and long overdue.

1.2 Direction and Scope

The continuing problem and enduring challenge is to predict materials failure behavior in general situations based upon knowledge of the loads and minimal failure data obtained from testing. The data that are usually found in comprehensive handbooks and in textbooks are those of the uniaxial tensile and compressive strengths. The historic activity for isotropic materials has always been and still is to attempt to theoretically prescribe the general, multi-dimensional envelope that discriminates between a safe and stable state of stress, and that of the total loss of the load-bearing capability. Ideally, this failure envelope or failure criterion

would be calibrated by only two or three data points obtained from the material of interest. That this could even be possible is at first thought quite surprising, yet it has been enormously successful in some related situations.

The general linear, elastic deformation (before failure) for all homogeneous and isotropic materials is predictable from only two mechanical properties. Even in the failure domain, the failure (yielding) of very ductile metals is convincingly predictable from the Mises criterion, which is a one-property form. The drive then is to extend this capability to all materials types, and to do so by a method requiring only a minimal number of material-specific properties that are readily accessible from testing. A general theory of failure behavior that is the counterpart and companion piece to the theory of elasticity (and plasticity) is needed, and it must be in a formalism that is suitable for broad application. This general failure theory, with all of its implications and many of its applications, is the specific direction of this book.

It would be a reasonable question to ask why the subject of physically based failure criteria for homogeneous materials seems to be so difficult, and why it is so late to come to fruition. The short answer is that failure is the ultimate non-linearity. It has come to be recognized as one of the "all time" difficult problems or disciplines. Why then, after so many unsuccessful attempts, might it now yield to rational and reasonable solution? Again, the short answer and probable reason is that it could not have been developed until after the advent of fracture mechanics, which itself is of relatively recent occurrence. Fracture mechanics opened the door to new and deeper ways of characterizing failure. The same can be said of the more recent developments for plasticity theory.

The heart of the book is the development of failure theory for homogeneous and isotropic materials—the perennial problem. The explicit and full derivation is given in one high-concentration chapter: Chapter 4, "The Failure Theory for Isotropic Materials." Everything that comes before the derivation involves the necessary steps leading up to it, and most of what comes after it keys off from it to give a comprehensive account of the subject. The derivation itself represents the synthesis and completion of the long-sought theory of materials failure.

The scope of applicability of the failure theory to be treated is unusually broad. It is for all full-density, homogeneous materials in macroscopic states of imposed stress. As such it is applicable to metals, polymers,

ceramics, glasses, and some geological materials. Also, special classes of highly anisotropic materials (carbon fiber composites) will be included. Because the failure theory is calibrated by a minimal number of standard, intrinsic strength properties of the material, it has great flexibility in application. Essentially, all standard engineering materials come within its field of view and use. Materials not covered by the theory are cellular materials (foams), granular materials, and other inhomogeneous materials forms. Their failure criteria require separate development.

1.3 Guides for Utility

A scan of the Contents in the preliminary pages of this book shows a very wide range of topics, yet they all are related to the main theme of materials failure. The book effectively starts with an historical summary in Chapter 2, and finishes with the probabilistic treatment of failure in the final Chapter 16. In between, the chapters have an organization that follows logical steps of development, with each new chapter building upon the previous ones. Ideally, the entire book would be approached in this exact sequential manner. To place the steps of development into some perspective, it would be hard to imagine constructing a successful failure criterion for anisotropic fiber composites without first gaining a thorough understanding of isotropic materials. Isotropic materials failure must come first, and further complexity builds up from there.

A special feature of the coverage is a thorough treatment of ductile-versus-brittle types of failure behavior. If especially difficult topics such as this were not covered here there would not be a credible case for a general treatment of failure. Ultimately this inclusion should help deepen the usefulness and significance of the book.

Since this is a macroscopic failure theory for homogeneous materials, specific failure mechanisms are not explicitly involved. Thus, such failure mechanisms as kink bands in aligned fiber composites and void nucleation in metals and other common mechanisms, each of which must be treated individually and are very helpful in that way, are nevertheless not needed here. The complete array of such mechanisms are implicit in the overall failure criteria. For a particular materials type in a particular stress state, the controlling failure criterion represents the underlying and dominant failure mechanism at that condition.

This book is related to the website www.failurecriteria.com [1.4], which was established in 2008 and has become quite widely utilized. Although the book and the website concern the same general topic, they differ greatly in organization, have different purposes, and include some significant differences in content. The book conforms to academic standards of scholarship, and in effect it formalizes the subject matter of the website.

References

[1.1] Timoshenko, S. P. and Goodier, J. N. (1970). *Theory of Elasticity*, 3rd ed., McGraw Hill, New York.

[1.2] Sokolnikoff, I. S. (1983). *Mathematical Theory of Elasticity*, 2nd ed., Krieger, Malabar, FL.

[1.3] Hill, R. (1998). *The Mathematical Theory of Plasticity*, Oxford University Press, Oxford.

[1.4] Christensen, R. M., (2013). "Failure Theory for Materials Science and Engineering," http://www.failurecriteria.com/.

2
History, Conditions, and Requirements

Much of physics has always been concerned with identifying basic effects, then formulating their behaviors as general theories, and finally establishing the limits of validity of the ensuing mathematical idealizations. The task was or is never easy, but at least it is much less complicated when the physical quantity of interest can be taken as a scalar or even a vector rather than as a tensor. For example, heat transfer and radiation propagation fall within the former category. On the other hand, describing the load-bearing capability of materials necessarily involves stress, which brings in second- and fourth-order tensors.

Stress (and strain) is a second-order tensor involving nine individual and independent entities. Symmetry of the shear stress terms reduces the number to six independencies. But when one begins to consider the limits of behavior in general and the limits of the stress tensor in any particular material, then one must distinguish between positive and negative normal stresses (non-shear stresses) and consider all the possible combinations of all these different forms. Suddenly the number of possible combinations of contributing terms becomes unlimited. The overall scale of the problem becomes almost unthinkably large, at least in terms of the raw elements of the problem.

It is probably fortunate that those involved in the historical search for the controlling failure criterion of materials did not perceive the problem in such stark and formidable terms. Quite understandably there was considerable optimism that the existing very high levels of physical insight combined with some ingeniously clever testing would uncover the treasure of a simple but universal failure criterion. Many of history's greatest scientists took part in the search. This is the account of the historic quest.

2.1 Historical Review

It is technically helpful and surprisingly revealing to examine the history for this field of materials failure. A brief account will now be given of the major historical turning points, since they are of importance in assessing the current state of knowledge and practice in this field. From this point on, in the historical account and in the work to follow, the term "failure" or "failure characterization" will be used in an inclusive sense as applying to either or both of the conditions of yielding or actual failure. All considerations here are with homogeneous and isotropic materials, where the historical goal was to find a general failure criterion that would cover the full range from ductile to brittle types of materials.

Probably the most misunderstood but definitely the most prominent of all failure concepts is that of the Coulomb–Mohr type. Its antecedents go back to nearly the beginning of all of mechanics, and its popular reception continues up to the present day. It comes from the original conception of Coulomb [2.1] that on a failure surface the shear stress at failure τ is related to the normal stress σ acting across the failure surface as

$$\tau \leq c - \mu\sigma$$

where the two parameters c and μ are material-specific. With $\mu = 0$ this is just the maximum shear stress criterion usually known today as the Tresca form.

Coulomb was a brilliant engineer working at military installations in early adulthood, and much concerned with structures and stability—possibly of soil embankments. In fact, he labeled parameter μ as the coefficient of internal friction, suggestive of granular material flow or incipient flow. This simplest form relates shear stress on an assumed failure plane to the transverse normal stress (tensile or compressive). Thus, the failure criterion is assumed to depend only upon the tractions acting upon the failure surface, and it is independent of the stress components in the plane of the assumed failure surface. The two parameters, c and μ, determine the intercepts and the slopes of the straight-line failure envelopes in Fig. 2.1.

Mohr [2.2] recognized that the straight-line envelopes shown in Fig. 2.1 would not be likely, and that the envelopes in general would be curved,

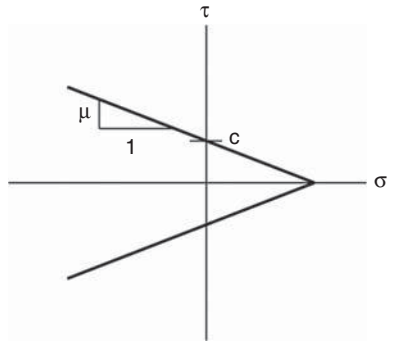


Fig. 2.1 *Coulomb failure criterion.*

concave inward. But in the interests of a usable form, Mohr also gave an interpretation especially applicable to Coulomb's two-parameter criterion. The maximum and minimum principal stresses are used to form a "Mohr's circle" construction within the linear failure envelopes of Fig. 2.1. Incipient failure occurs when the circle has tangency with the envelopes, as can be illustrated for uniaxial tension and compression. This procedure will be outlined in Chapter 3.

With the slopes of the envelopes in Fig. 2.1 interpreted as the friction coefficient, then the other parameter is the cohesive strength. Alternatively, the uniaxial tensile and compressive failure stresses can be used to calibrate the two failure parameters. This criterion and procedure is usually now known as the Mohr–Coulomb or Coulomb–Mohr form.

For brittle materials the Coulomb–Mohr method predicts that uniaxial and equibiaxial tensile failure are at the same levels, but that equitriaxial tensile failure is much stronger. This latter unphysical behavior has no supporting evidence, and probably occurs only in the ductile range and certainly not in the brittle range of behavior.

In testing geological materials, von Karman [2.3] and also Böker [2.4] showed that the Coulomb–Mohr criterion is only partially successful when the stresses are somewhat compressive, and not at all successful otherwise, either in the very compressive or slightly compressive region, the latter of which is unfortunately the most limiting region. This work was carried out with support and guidance from Prandtl. Voigt also found that the Coulomb–Mohr form could not model general failure behavior, but he was so negative about its prospects that he left his research unpublished (Timoshenko [2.5]).

There is nothing particularly fundamental about the Coulomb–Mohr approach, as is often claimed. It is simply a well-developed hypothesis about failure, but it is not unique, and not necessarily useful unless proven to be so. Examples will be given later, further showing its deficiencies. None of this detracts from the original contribution of Coulomb, nor of Mohr for that matter. Coulomb's recognition that the shear stress at failure would be influenced by the transverse normal stress is absolutely correct, and his form is of considerable application in soil mechanics and rock mechanics over a limited range. The useful application of the Coulomb–Mohr type to materials science and technology has been embryonic, and has never developed beyond that stage.

It was Maxwell who had the acuity and perception to first see the use of energy as a possible or partial failure criterion. In correspondence with Lord Kelvin in 1856, Maxwell anticipated what is now known as the Mises criterion by saying "I have strong reasons for believing that when the strain energy of distortion reaches a certain limit then the element will begin to give way" (Timoshenko [2.5]). Much later, Huber [2.6], and still later Mises [2.7], gave this the form in which it is used today for the yielding of ductile metals.

Over the years the apparent appeal of various forms of energy as a general failure criterion has been elusively strong. Beltrami [2.8] proposed that the total energy—energy of distortion plus energy of volume change—be used as the general criterion, but this is clearly contradicted by the ductile metals case, and other materials as well. Only the Mises case for distortional energy is indisputably correct when applied to ductile metals, but not for any other type of materials. Taylor and Quiney [2.9] produced the conclusive experimental verification for this ductile-metals criterion.

In addition to other approaches, maximum normal stress and maximum normal strain failure criteria had distinguished and strong adherents—Rankine and Lamé for the former, and Saint-Venant and Poncelet for the latter (Timoshenko [2.5]). Neither came into a state of explicit validation and meaningful application, although they certainly added to the historical state of confusion. These two criteria are sometimes still used, with no apparent general justification. The maximum shear stress criterion was judged by Tresca [2.10] to be the proper form for ductile metals. This would later be re-examined and re-evaluated by Taylor and Quiney, as noted above.

It should not be inferred that the historical efforts were not extremely helpful. As already implied, the profound deductions of Coulomb and of Maxwell were the most important steps of all time in this field. Perhaps each was motivated by a particular materials behavior at opposite ends of the scale. Other scientists certainly understood the essential ingredients of the problem. For example, as discussed by Todhunter and Pearson (1893), Lord Kelvin held the view that "It may be inconceivable that any amount of uniform pressure applied to the surface of a solid sphere of isotropic material would cause it to rupture, but it is also very difficult to believe that a uniform tension, if it could be applied to its surface, would not, were it indefinitely increased, produce rupture" (Thomson and Tait [2.11]). However, that understanding did not lead to progress toward a failure criterion for application. Despite more than a century of intense inquiry, Love [2.12] would write: "The conditions of rupture are but vaguely understood." The term "rupture" had about the same implied meaning as "failure" or "fracture" would today.

Another two-parameter failure criterion was given much later by Drucker and Prager [2.13]. The Mises criterion is seen as a cylindrical surface in principal stress space. The Drucker–Prager form generalized the Mises criterion to take the form of a cone in principal stress space. Although both the Coulomb–Mohr and Drucker–Prager forms are widely used in soil mechanics, both failed to gain use for general engineering materials. Both criteria will be discussed further in Chapter 3.

After the many attempts to cast the Coulomb–Mohr hypothesis into a generally realistic, accessible form never came to fruition, and the quest for a general energy criterion necessarily came to an unsuccessful end, a degree of pessimism seemed to influence the further efforts to find a general criterion. Thereafter, the attempts were direct postulations of particular forms, appealing on some basis to each originator, but apparently done with a motivation to just see what would happen. None of the general approaches in the modem era had a solid basis in a physical derivation along with a critical examination; there have been only demonstrations. For some examples, see the survey by Paul [2.14].

The consideration of the Coulomb–Mohr criterion had extended over more than a hundred years and still had not produced the desired result. The new phase of the search mentioned above centered around a consensus that two parameters could not do the job, and more—perhaps

many more—than two parameters would be needed. A consequence of the many-parameters approach was that emphasis would be given to applications to particular classes of materials with no expectation of generality. Some would be for metals, some for polymers, some for rocks, and so on. Such empirical forms could certainly have considerable utility, but would have undefined and vague limits of applicability. For example, the book by Jaeger, Cook, and Zimmerman [2.15] has a very good summary of forms applicable to rock mechanics. The book by Yu [2.16] has a three-parameter generalization of the Coulomb–Mohr criterion, but still includes most of its deficiencies. Individual multi-parameter failure criteria are typified by those of

Willam and Warnke
Bresler and Pister
Zienkiewicz and Pande.

All these and many more can be accessed through the Internet. A fairly complete list of the many different failure models through 2004 is given by Yu [2.16]. None of these have gained wide usability and acceptance.

The only related and truly significant development in modern times has been the remarkable emergence of fracture mechanics. It has indeed been a tremendous development with far-reaching implications. Griffith [2.17] is widely and justifiably credited with originating fracture mechanics. In lesser-known work, Griffith [2.18] also went a step further and adapted the crack-induced failure effect into homogeneous material behavior, and thereby prescribed a failure criterion with a universal ratio of uniaxial compressive-to-tensile failure stresses at exactly the value of 8. Murrell [2.19] extended this Griffith treatment and derived the compressive-to-tensile ratio as being exactly 12. Neither result has been found to have general applicability, nor could they. The relationship, if any, of general criteria of the type sought here to fracture mechanics has remained clouded in this more modem era. Much more will be said later in this book of the relationship between fracture behavior and failure criteria for homogeneous materials.

Arriving at nearly present times, clearly the state of general three-dimensional failure characterization for homogeneous and isotropic materials is not satisfactory. There have been literally hundreds and perhaps even thousands of failure forms displayed and proposed for use; but still there is no generality, only empiricism.

Despite the long, difficult, and unpromising history, there remains the possibility that the expanding knowledge base of the modern era may provide the advantage needed for a more successful formulation of failure theory.

2.2 Conditions and Requirements of Study

Well-constructed failure theories can discriminate safe states of stress in materials from states of certain failure, based upon calibration by a minimal number of failure-type mechanical properties. The specific purpose here is to provide failure criteria for general types of materials. Two of the conditions that are taken to apply are those of a macroscopic scale of consideration and the corresponding macroscopic homogeneity of the material.

The concepts of macroscopic scale and macroscopic homogeneity have connotations familiar to everyone. However, trying to define these concepts in absolute terms is extremely difficult. Macroscopic homogeneity is taken to be the condition that the material's constitution is the same at all locations. Thus the problem is shifted to the precise meaning of the term "location," which depends upon the scale of observation. Suffice to say, the scale of observation is taken such that all the common forms of materials are included, such as metals, polymers, ceramics, glasses, and some geological materials. Materials which are excluded are porous materials, whether cellular or not, as well as granular materials.

In the case of metals, the macroscopic scale must be taken as being much larger than the size of the individual grains—presumably about an order of magnitude larger. Thus the macroscopic scale depends upon the type of material. For polymers it must be much larger than the characteristic molecular dimension. Other cases are similarly apparent. Effects at the nanometer and micrometer scales do not explicitly enter the macroscopic formulation. To attempt to bring in these effects would add parameters restricted to a particular materials type and obscure the clarity of a self-contained and self-consistent treatment at the macroscopic scale. However, after obtaining the general macroscopic treatment it is very interesting to compare and interpret with effects at the smaller scales for specific materials types.

Failure itself can sometimes be difficult to identify and even more difficult to define. For cases considered here, failure is taken to involve the

interruption of the usual linear, reversible range of behavior by a major change to irreversibility. Failure implies the material's lack of ability to sustain and support significant loads. In subsequent chapters the term "damage" will be defined as a subset of failure.

Failure in solids bears an interesting relationship to turbulence in fluids. Fully turbulent flow in fluids and the failure occurrence in solids mark the departure from linear control to a completely non-linear behavior. It does not appear possible to establish an analogy between turbulent flow and failure in generalized continua; nevertheless, both cases conform to the dominance by an ultimate non-linearity. In effect, both cases represent completion of the account of behavior prescribed by the balance laws and constitutive equations.

Some further perspective may be appropriate concerning the effort to theoretically characterize the failure of materials of various types through mathematical criteria. Perhaps no other scientific quest has had as much energy expended upon it historically with so little to show for the effort. The viewpoint here holds that the scientific advances of the modern era lay out the proper tools and directions which can finally bring failure criteria to a level of unification (and turbulence modeling as well). The insight offered by fracture mechanics is especially helpful in understanding brittle failure in homogeneous materials.

Several qualifications must be included concerning the purpose and scope of this effort. This book is not aimed as a literature survey; the goal here is to be technically discriminating rather than inclusive. The field of materials failure in general and failure criteria in particular is enormously broad, and a complete treatment of such matters would be unrealistic and probably impossible. The focus here is upon the theoretical basis for the specific failure criteria under consideration. The term "theoretical basis" signifies the physical and mathematical underpinnings of the various failure forms. No attention will be given to empirical forms that may seem to work well in certain situations but have no foundation from which to judge their probable generality. Although it is not intended to integrate broad datasets for many materials, which itself is a large and important but self-contained topic, guides and references to such sources will be given wherever possible. Critical datasets will be used to evaluate key theoretical developments.

Finally, most failure criteria to be reviewed are taken directly from (or intimately related to) results already presented in peer-reviewed

archived journals. Such publications have received critical examination and evaluation. While this certainly does not provide any kind of warranty, it does help to assure fidelity with important aspects of physical behavior, which is the central objective. Again, this book is not directed toward being a literature survey of the field, as that is best conducted through Internet search services. However, all works that are explicitly used here are fully referenced.

References

[2.1] Coulomb, C. A. (1773). In *Mémoires de mathématiques et de physique présentés à l'Académie royal des sciences par divers savants*, **7**, 343–82.

[2.2] Mohr, O. (1900). "Welche Umstände bedingen die Elastizitätsgrenze und den Bruch eines Materials," *Zeitschrift des Vereins Deutscher Ingenieure*, **44**, 1524–30.

[2.3] von Kármán, T. V. (1912). "Festigkeitsversuche unter allseitigem Druck," *Mitteilungen Forschungsarbeit Gebiete Ingenieurs*, **118**, 27–68.

[2.4] Böker, R. (1915). "Die Mechanik der Bleibenden Formanderung in Kristallinisch Aufgebauten Körpern," *Mitteilungen Forschungsarbeit auf dem Gebiete Ingenieurwesens*, **24**, 1–51.

[2.5] Timoshenko, S. P. (1953). *History of Strength of Materials*, McGraw-Hill, New York.

[2.6] Huber, M. T. (1904). "Przyczynek do Podstaw Wytorymabsci," *Czapismo Technizne*, **15**, 81.

[2.7] von Mises, R. (1913). "Mechanik des festen Körper im plastisch-deformablen Zustand," *Nachr. Gess. Wiss. zu Göttingen, Math.-Physik Klasse*, **1**, 582–92.

[2.8] Beltrami, E. (1889). "Considerazione Idrodinamiche," *Rend. 1 Lombardo Sci. Lettere*, **22**, 121–30.

[2.9] Taylor, G. I. and Quiney, H. (1931). "The Plastic Distortion of Metals," *Phil. Trans. Roy. Soc. London*, **A230**, 323–62.

[2.10] Tresca H. (1868). "Memoire sur l'ecoulement des corps solides," *Mem. Pres. Par. Div. Sav.*, **18**, 733–99.

[2.11] Thomson, W. (Lord Kelvin) and Tait, P. G. (1879). *Treatise on Natural Philosophy*, Cambridge University Press, Cambridge.

[2.12] Love, A. E. H. (1927). *A Treatise on the Mathematical Theory of Elasticity*, Cambridge University Press, Cambridge.

[2.13] Drucker, D. C. and Prager, W. (1952). "Soil Mechanics and Plastic Analysis or Limit Design," *Quart. Appl. Math.*, **10**, 157–65.

[2.14] Paul, B. (1968). "Macroscopic Criteria for Plastic Flow and Brittle Fracture," in *Fracture* (ed. H. Liebowitz), **II**, 313–496, Academic Press, New York.

[2.15] Jaeger, J., Cook, N. G., and Zimmerman, R. (2007). *Fundamentals of Rock Mechanics*, Wiley, New York.

[2.16] Yu, M. H. (2004). *Unified Strength Theory and its Applications*, Springer, Berlin.

[2.17] Griffith, A. A. (1921). "The Phenomena of Rupture and Flow in Solids," *Phil. Trans. Roy. Soc. London*, **A221**, 163–98.

[2.18] Griffith, A. A. (1924). "The Theory of Rupture," *Pro. 1st Int. Congr. Appl. Mech.*, 55–63.

[2.19] Murrell, S. A. F. (1963). "A Criterion for Brittle Fracture of Rocks and Concrete under Triaxial Stress and the Effect of Pore Pressure on the Criterion," *Pro. Fifth Symp. on Rock Mechanics*, in *Rock Mechanics* (ed. C. Fairhurst), Pergamon, Oxford, 563–77.

3
Isotropic Baselines

This chapter is concerned with the more explicit preliminaries that must be considered before full immersion in the complete failure problem. First is a brief discussion of failure criteria goals and objectives followed by the necessary commitment to either stress or to strain for describing the failure event. After that, the two widely used and historic one-property forms are examined: the Mises criterion and the Tresca criterion. Finally, the two historic two-property forms will be specified: the Coulomb–Mohr criterion and the Drucker–Prager criterion. The difficulties and limitations with all four of these forms will be examined. All considerations here concern isotropic materials.

3.1 Failure Characterization

The quasi-static failure of cast iron under torsion has the failure surface as inclined at 45 degrees to the longitudinal axis. Thus the failure surface takes the orientation that produces the maximum possible tensile stress component (principal stress) as acting upon it. A failure criterion must respect and reflect this physical characteristic for this particular materials type, but if it is to apply generally it must also be compatible with many other physical failure characteristics for many other isotropic materials types.

It is highly advisable that the failure characterization for any one materials type—epoxy polymers for example—follow directly as a specialization from a comprehensive form applicable to all homogeneous and isotropic materials. Success across the spectrum of all materials types, from very ductile to very brittle, would be much more likely to assure success for any one particular class of interest.

In the present work it will be required first that the general characterization admit the proper one-property criterion at one extreme, giving the correct results for ductile metals. Then at the other end of the

scale, realistic specializations must occur for the cases of brittle ceramics, glasses, and geological materials, which surely would require more than one property.

It is furthermore required that the general theory be calibrated by only a few (preferably two) failure properties. If many properties appeared to be needed, the situation would inevitably degenerate to being just that of adjusting or fitting parameters rather than that of an evaluation from mechanical properties. The same unsatisfactory situation would usually occur in merely fitting data for a single materials type with no regard for anything else.

It can be shown that the two-property general theory expounded in Chapter 4 predicts physical results for cast iron that are completely consistent with those described in the first paragraph above. On a scale going from limit case ductility to extreme brittleness, this example of cast iron corresponds to a moderately brittle case and thus is intermediate on that scale. Other examples from other major classes of materials are equally important to examine.

The mainstream topics of traditional interest have been composed mostly of isotropic materials. However, anisotropic materials in general and fiber composite materials in particular are of great contemporary interest and relevance. These broader classes of materials will also be included. Beyond the idealization of quasi-static failure lie the even more difficult classes of fatigue and rate-dependent failure. These too should be and will be covered. Nevertheless, the priority order taken here is that a consistent and self-contained formulation of quasi-static, isotropic material failure must come first. If this most basic aspect of failure cannot be correctly treated, then there is little likelihood of succeeding with the even more difficult problems.

3.2 Stress versus Strain

Probably the most longstanding issue with failure criteria is whether they should be expressed in terms of stresses or in terms of strains. The discussion goes back to the very beginning and continues even now. For example, some aerospace organizations specify design procedures in terms of stress, while some others use strain. Typical technical exchanges on the subject remain inconclusive. Still, the uses of failure criteria are manifold, and everyone is left with the problem of making the best possible choice.

It is likely that the stress-versus-strain issue will remain as a strongly held division between the practitioners for some time to come. Experience often involves trade-offs, contradictions, and anomalies. However, at the research base there should be a logical resolution of the problem.

The most commonly used form is that of the Mises criterion expressed in terms of stress. Other criteria also in common usage are the maximum shear stress criterion (Tresca), the maximum normal stress criterion, the maximum normal strain criterion, and dilatational criteria in terms of either stress or strain. The Mises criterion applies to isotropic materials, and therein lies an unforeseen complication. The stress–strain relations for isotropy can be used to convert the Mises criterion in terms of stresses into a criterion expressed in terms of strains. A surprising thing happens: the strain form is the same as the stress form. In other words, the Mises criterion is form-invariant between stresses and strains.

It is a short step from the above incontrovertible fact to assume that all criteria can be switched from stress to strain (and vice versa) with no basic change in character. Such is not the case however, especially when considering the broader classes of anisotropic forms, but also for some isotropic cases. As an example, the maximum normal stress criterion and the maximum normal strain criterion are of decidedly different forms when interconverted.

Another fact to be considered is that the Mises criterion is applicable only to ductile metals and is in serious error for all other materials types. These matters signal the difficulties that ensue in trying to go from stress to strain or the reverse in considering failure. All these complications trace back to the tensor-valued nature of stress and strain. If they were scalars the difference would be trivial. A rational basis is needed for making the judgment call on the stress-versus-strain uncertainty.

In this book, and in the supporting peer-reviewed papers upon which it is based, stress is taken as the fundamental form to be used with failure criteria. This is the direct consequence of the following conditions. Stress must be used if one wishes to have compatibility with fracture mechanics in the brittle range and with dislocation dynamics in the ductile range. Both of these classical theories require formulations in terms of stress. To not have union with these two anchor points of physical reality would impose major difficulties, practically as well as conceptually.

In a totally different direction, a very interesting failure effect is exhibited by materials with time dependence (even an almost undetectable

degree of time dependence, as with polymers below the glass transition temperature). For the application of constant load (stress) the situation is termed as creep. If the load is maintained long enough—even years in some cases—at some point total failure can occur: creep rupture. Conversely, when a state of time constant strain is applied, termed as relaxation, no failure occurs no matter how long the strain state is maintained. Even the technical term "relaxation" is descriptive of what is happening at the molecular scale in the case of constant strain. Clearly, in these problems, stress rather than strain is the stimulus for failure.

Yet another indication of the more sound basis for stress is that it is applicable both to solid materials and to fluids. Strain (deformation) is not, as it has no meaning for fluids. Some viscoelastic fluids can and do fail, and stress necessarily must be used in such cases.

Stress, not strain, will be used here in characterizing failure criteria for the physical reasons just stated. This decision will be further substantiated by the success of stress-based forms and the ambiguity of strain-based forms in particular examples.

3.3 Mises and Tresca Failure Criteria

All books on the mechanics of materials have sections on strength and failure criteria. Usually the Mises and Tresca criteria are presented jointly with little discrimination or recommendation between them. What is more, often little else in the way of failure criteria is presented, and the non-expert reader can be left with the impression that these two criteria cover all or nearly all materials and that it does not greatly matter which one is used. Some books even grant these two criteria the status of being classical results. A general question to be posed here is this. Can either or both of these criteria be considered to be classical results, or are they merely longstanding comfortable forms but of no special distinction?

In this section a careful look is taken of the two criteria, Mises and Tresca, mainly in comparison with each other to see if there should be some inherent preference, even at this level. Concurrently, some observations will be presented of how each fits into the larger picture involving generality beyond just that of the usual ductile metals.

The main interpretation of the Mises criterion is that it represents a critical value of the distortional energy stored in the isotropic material, while the Tresca criterion is that of a critical value of the maximum

shear stress in the isotropic material. Both criteria are widely discussed in many places. Historically, the Tresca form was considered to be the more fundamental of the two, but the Mises form was seen as an appealing, mathematically convenient approximation to it. Now, both are usually stated side by side with little or no preference.

The two criteria are specified below in principal stress space. Both are one-parameter forms, specified by either the uniaxial tensile strength, T (with $C = T$), or the shear strength, S.

Mises Criterion, Critical Distortional Energy

$$\frac{1}{2}\left[(\sigma_1 - \sigma_2)^2 + (\sigma_2 - \sigma_3)^2 + (\sigma_3 - \sigma_1)^2\right] \leq T^2 \tag{3.1}$$

where

$$C = T$$
$$S = \frac{T}{\sqrt{3}}$$

Tresca Criterion, Critical Shear Stress

For the principal stresses ordered as $\sigma_1 \geq \sigma_2 \geq \sigma_3$, then

$$\frac{1}{2}(\sigma_1 - \sigma_3) \leq S \tag{3.2}$$

For the principal stresses not ordered,

$$\begin{aligned} \frac{1}{4}(\sigma_1 - \sigma_2)^2 &\leq S^2 \\ \frac{1}{4}(\sigma_2 - \sigma_3)^2 &\leq S^2 \\ \frac{1}{4}(\sigma_3 - \sigma_1)^2 &\leq S^2 \end{aligned} \tag{3.3}$$

where

$$C = T$$
$$S = \frac{T}{2}$$

The three separate forms in (3.3) are for the maximum shear stresses in the three principal planes.

Both of these single-property criteria can be calibrated on either T or S. Figs. 3.1 and 3.2, for biaxial stress states, calibrate both criteria on T and then on S.

Case I

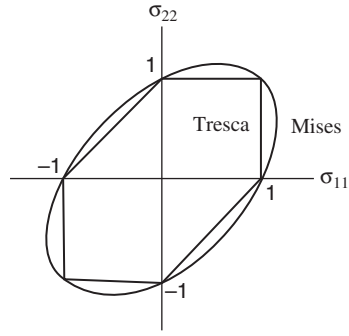


Fig. 3.1 *Calibration on T.*

Case II

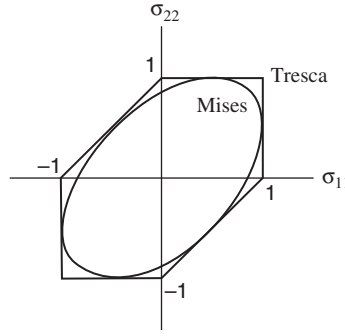


Fig. 3.2 *Calibration on S.*

The maximum difference between the Mises and Tresca forms for both Cases I and II is 14.4%. The essential and striking differences between the two are the corners that occur in the Tresca form and their complete absence in the Mises form.

Two basic questions are immediately apparent:

Which procedure is better grounded: calibrating on T or calibrating on S?

Which criterion is more physically relevant: Mises or Tresca?

The steps of reasoning used here to answer these two questions are as follows. Calibrating on S might seem more logical because yielding is caused by distortional not dilatational states in ductile materials, and distortion more nearly and easily associates with shear stress rather than with uniaxial tensile (or compressive) stress.

However, shear stress involves only one parameter, and that does not naturally generalize to a two-parameter form which is necessary for non-perfectly ductile materials. On the other hand, $T = C$ naturally and automatically generalizes to T and C ($T \neq C$) in the more general cases. So T and C, not S, is the natural combination that converges to $T = C$ for a one-parameter form. Case II, calibrating on S, is accordingly rejected.

With calibration by T, giving Case I, there is still the choice between the Mises and Tresca forms. Mises is smooth, while Tresca has corners. At the crystal level (single grain) yielding does associate with dislocation movement on slip planes. This is caused by shear stress on the slip system (resolved shear stress). It would be tempting to say that this justifies and validates the Tresca criterion. But that is only part of the consideration. The condition of isotropy implies and applies to polycrystalline aggregates with the individual crystals taking all possible orientations. Dislocation-induced plastic flow occurs over many slip systems. Furthermore, dislocation pile-up occurs at the grain boundaries. The vastly more complex behavior at the aggregate level compared with the crystal level must involve averaging over a wide variety of physical conditions and effects. This averaging has a smoothing effect that is much more supportive of the smooth Mises criterion than of the non-smooth Tresca form.

These conclusions support and justify using the Mises (not Tresca) criterion for very ductile materials, and calibrating on T (not S).

Further relevant background considerations are as follows. Tresca did do testing of metals that at the time seemed to support the maximum shear stress criterion. But that testing was superseded by the later excruciatingly careful testing performed by Taylor and Quinney, as shown in Chapter 6. The Taylor–Quinney results support the Mises criterion.

All $T = C$ materials are metallic, polycrystalline aggregates. It is necessary to carry out orientation averaging to determine the effective isotropic moduli properties. It is a parallel and consistent circumstance

that orientational averaging must also be carried out to determine the strength for polycrystalline aggregates. Mises (3.1) is consistent with this, while Tresca (3.3) is not. In fact, Mises (3.1) is a composition or a type of average of the three separate criteria in (3.3), Tresca.

The non-smooth behavior evinced by the Tresca criterion usually associates with the competition of failure modes such as with a ductile flow mode and a brittle fracture mode. But those competitive effects are not present with ductile materials. Now consider a true two-dimensional type of continuum to see how it behaves. This is not plane stress or plane strain that are still three-dimensional behaviors. In this truly two-dimensional case it is found that a maximum shear stress criterion (Tresca) and a maximum distortional energy criterion (Mises) are identical, both giving smooth behaviors with continuous first derivatives Then, in going to three dimensions the Mises form continues this smooth behavior but the Tresca form brings in corners. The Tresca behavior in three dimensions is an artifact of describing the maximum shear stresses in the three principal coordinate planes. This ignores the effects that occur at smaller scales in polycrystalline aggregates, and the averaging necessary to reach macroscopic behavior.

Even though the maximum difference between the Mises and Tresca criteria is only about 15% this difference represents a systemic error (divergence) on the part of the Tresca criterion, and it should not be used for any isotropic materials, even for ductile metals. It is inappropriate to place the Tresca criterion on the same level as the Mises criterion, as is done in most tutorial works. The Tresca criterion merits only an historical reference. This is consistent with the fact that the Tresca criterion is the limiting case of the two-parameter Coulomb–Mohr criterion which itself is only of historical interest. As discussed in Chapter 4, the Mises criterion is the limiting case of a viable, completely general, modern failure criterion.

3.4 Drucker–Prager Failure Criterion

The Drucker–Prager [3.1] theory was probably motivated by the desire for a general theory that would reduce to the Mises form in the limit, rather than the Tresca form. A course of action to accomplish this is quite direct. Simply replace the Coulomb–Mohr six-sided pyramid in principal stress space (to be described next) by a circular cone.

Following this direction, write the possible failure criterion as

$$a\sigma_{ii} + b\sqrt{s_{ij}\,s_{ij}} \leq 1 \tag{3.4}$$

where s_{ij} is the deviatoric stress tensor involved in a Mises criterion, but now the term a explicitly brings in the mean normal stress effect which is implicit in the Coulomb–Mohr formalism.

Parameters a and b are to be determined such that (3.4) predicts uniaxial stress failure at the tensile value T and the compressive value C. It is easily shown that this gives (3.4) as

$$\frac{1}{2}\left(\frac{1}{T} - \frac{1}{C}\right)\sigma_{ii} + \frac{1}{2}\left(\frac{1}{T} + \frac{1}{C}\right)\sqrt{\frac{3}{2}s_{ij}\,s_{ij}} \leq 1 \tag{3.5}$$

Finally, writing (3.5) in terms of components gives

$$\begin{aligned}&\frac{1}{2}\left(\frac{1}{T} - \frac{1}{C}\right)(\sigma_{11} + \sigma_{22} + \sigma_{33})\\ &\quad + \frac{1}{2}\left(\frac{1}{T} + \frac{1}{C}\right)\left\{\frac{1}{2}\left[(\sigma_{11} - \sigma_{22})^2 + (\sigma_{22} - \sigma_{33})^2 + (\sigma_{33} - \sigma_{11})^2\right]\right.\\ &\qquad \left. + 3\left(\sigma_{12}^2 + \sigma_{23}^2 + \sigma_{31}^2\right)\right\}^{\frac{1}{2}} \leq 1\end{aligned} \tag{3.6}$$

As shown, only the positive sign is associated with the square roots in (3.4)–(3.6).

In principal stress space, (3.6) is that of a conical failure surface. The symmetry axis of the cone makes equal angles with the three principal axes. When $T = C$ the failure surface becomes the circular cylindrical form of Mises.

A simple example shows the inappropriateness of the Drucker–Prager criterion for representing materials failure. Take for example the case of cast iron, which has a $T/C = 1/3$ value. Next specify the condition of equal biaxial compressive stresses

$$\begin{aligned}\sigma_1 &= \sigma_2 = -\sigma\\ \sigma_3 &= 0\end{aligned} \tag{3.7}$$

Using these specifications in (3.5) gives

$$\left(-1+3\frac{T}{C}\right)\sigma \leq 2T \tag{3.8}$$

thus

$$\frac{T}{C} \leq \frac{1}{3} \tag{3.9}$$

allows unlimited compressive stresses.

The Drucker–Prager criterion thus predicts that cast iron and all materials with $T/C \leq 1/3$ will support unlimited stress magnitudes in equal biaxial compression. This, of course, is impossible. Looking ahead to Chapter 4, the failure theory derived there predicts the failure stress in equal biaxial compression as $-1.55C$ for $T/C = 1/3$. Clearly, the Drucker–Prager criterion is not suitable as a failure criterion for engineering materials.

The general difficulties with the Drucker–Prager criterion, as well as with the following Coulomb–Mohr criterion, are discussed in Christensen [3.2]. Like the Coulomb–Mohr form, the Drucker–Prager form receives attention and usage, though not to as great an extent.

3.5 Coulomb–Mohr Failure Criterion

The procedure for the Coulomb–Mohr (or Mohr–Coulomb) theory uses the Mohr's circle construction as shown in Fig. 3.3, Mohr [3.3].

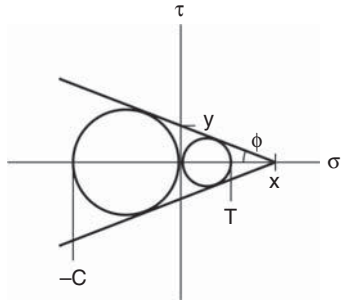


Fig. 3.3 *Coulomb–Mohr failure criterion characteristics.*

On the normal stress, σ, versus shear stress, τ, axes plot the circles for uniaxial tension and compression to failure, as shown in Fig. 3.3. Then two linear failure envelopes are taken as just being tangent to the two circles. In order to evaluate the safety for any given three-dimensional state of stress, the maximum and minimum principal stresses are used to form a circle on the plot. If this circle is inside the failure envelopes then there is no failure, otherwise there is failure of the material. For example, take a state of simple shear stress. Form a circle with its center at the origin in Fig. 3.3 and just tangent to the failure envelopes. This circle intersects the τ axis at the values of the shear stress at failure, S, given by

$$S = \frac{TC}{T + C} \tag{3.10}$$

where T and C are the failure levels in uniaxial tension and compression.

The general characteristics of the construction in Fig. 3.3 are given by

$$x = \frac{TC}{C - T}$$
$$y = \frac{1}{2}\sqrt{TC}$$

and

$$\sin\phi = \frac{C - T}{C + T} \tag{3.11}$$

The procedure just outlined for failure assessment from Fig. 3.3 can be expressed in analytical form. Let σ_1, σ_2, and σ_3 be the principal stresses, with σ_1 being the largest and σ_3 being the smallest in an algebraic sense. From relations (3.10) and (3.11) and Fig. 3.3, the failure criterion can be shown to be given by

$$\frac{1}{2}\left(\frac{1}{T} - \frac{1}{C}\right)(\sigma_1 + \sigma_3) + \frac{1}{2}\left(\frac{1}{T} + \frac{1}{C}\right)(\sigma_1 - \sigma_3) \leq 1 \tag{3.12}$$

The first term in (3.12) shows the normal stress effect, and the second term the maximum shear stress effect. Relation (3.12) directly reduces to

$$\frac{\sigma_1}{T} - \frac{\sigma_3}{C} \leq 1 \tag{3.13}$$

This extremely simple form constitutes the entire two-property Coulomb–Mohr failure criterion.

In the limit of $T = C$ the Coulomb–Mohr form reduces to the maximum shear stress criterion of Tresca. In principal stress space this is given by the infinite cylinder of hexagonal cross-section. When $T < C$ the Coulomb–Mohr form becomes a six-sided pyramid in principal stress space, and this six-sided pyramid has but three-fold symmetry rather than six-fold symmetry. An analytical form for this criterion in terms of invariants has been given by Schajer [3.4], although it is far easier to directly use the forms given here. The Coulomb–Mohr theory has been used in applications as diverse as nano-indentation and large-scale geophysics.

As with the Drucker–Prager example, it is not difficult to find realistic examples where the Coulomb–Mohr criterion produces blatantly incorrect results. Some examples are as follows. Take a three-dimensional compressive stress state specified by

$$\begin{aligned} \sigma_1 &= \sigma_2 = -\sigma \\ \sigma_3 &= -2\sigma \end{aligned} \tag{3.14}$$

Substituting (3.14) into (3.13) gives

$$\left(-1 + 2\frac{T}{C}\right)\sigma \leq T \tag{3.15}$$

Thus the Coulomb–Mohr criterion asserts that materials such as brittle polymers and a wide range of other materials can withstand unlimited compressive stresses in the stress state (3.14) when

$$\frac{T}{C} \leq \frac{1}{2} \tag{3.16}$$

Unfortunately this allows unlimited compressive stresses.

This prediction is completely and egregiously unrealistic. The stress state (3.14) can be thought of as an hydrostatic pressure of magnitude σ with a superimposed uniaxial compressive stress also of magnitude σ. Now, superimposed hydrostatic pressure does have a strengthening effect on any other stress state. From typical data, an hydrostatic pressure equal to twice the uniaxial compressive failure stress would be expected to approximately double the superimposed uniaxial compressive strength.

The failure theory from Chapter 4 applied to the stress state (3.14) with $T/C = 1/2$ gives the failure solution as $\sigma = 2.23$C, in agreement with the typical data expectation. The Coulomb–Mohr prediction of infinite strength in this stress state for these materials classes is totally unacceptable.

The Coulomb–Mohr form was very appealing in its simplicity of concept and its presumed generality and ease of use. There was great enthusiasm for it in the early 1900s when it came into general use. Then it was shown by von Karman [3.5] and Böker [3.6] to have considerable limitations for brittle materials—specifically, natural minerals. Much effort has been expended over the years to "correct" and generalize this most simple and direct form of the Coulomb–Mohr criterion, but to no special advantage or usefulness. It has continued to be used—perhaps because of the perceived lack of a suitable two-property alternative—but nevertheless the above examples unequivocally prove that the Coulomb–Mohr theory is fundamentally incorrect and should never be used, even though it is widely disseminated and codified.

3.6 The Bottom Line

Of the four standard and widely used failure criteria for isotropic materials,

Mises
Tresca
Drucker-Prager
Coulomb-Mohr

only the Mises criterion is of lasting significance. Although applicable only to very ductile metals, it is the classic result in that specific application, and it surely will endure. As shown, the other three criteria fail to establish the credentials necessary for permanency, and in fact can be misleading and patently incorrect in particular materials applications. Worst of all, none of the four can be used as general purpose failure criteria, as they are completely inappropriate to the task. Furthermore, it is probably obvious that adding another parameter or two to the Coulomb–Mohr or Drucker–Prager forms would be exceedingly unlikely to convert what is a completely unacceptable form into a transparently successful form. A totally new and fundamentally different approach is clearly required.

Although the results of this chapter do not lead to any useful general failure form(s), they do provide a hint of where to look for opportunity. In particular, the Coulomb–Mohr criterion is completely linear in the stress components and is unacceptable. Other linear and quasi-linear forms such as the Drucker–Prager criterion are also unacceptable for the same reason. It is probable that a general failure theory would involve some non-linear physical effects in some manner.

Going even further, it is still conceivable that there could be a reasonably straightforward failure descriptor, but it probably must have a non-linear formalism to support it. And if the Mises criterion gives any indication, this general non-linear formalism could be quadratic in nature. This possible opening will be pursued in Chapter 4.

References

[3.1] Drucker, D. C. and Prager, W. (1952). "Soil Mechanics and Plastic Analysis or Limit Design," *Quart. of Applied Mathematics*, **10**, 157–65.

[3.2] Christensen, R. M. (2006). "A Comparative Evaluation of Three Isotropic, Two Property Failure Theories," *J. Appl. Mech.*, **73**, 852–9.

[3.3] Mohr. O. (1900). "Welche Umstande Bedingen die Elastizitatsgrenze und den Bruch eins Materials," *Zeitschrift des Vereins Deutscher Ingenieure*, **44**, 1524–30.

[3.4] Schajer, G. S. (1998). "Mohr–Coulomb Failure Criterion Expressed in Terms of Stress Invariants," *J. Appl. Mech.*, **65**, pp. 1066–8.

[3.5] von Karman, T. V. (1912). "Festigkeitversuche unter allseitigem Druck," *Mitteilungen Forschungsarbeit Gebiete Ingenieurs*, **118**, 27–68.

[3.6] Böker, R. (1915). "Die Mechanik der Bleibenden Formanderung in Kristallinish Aufgebauten Korpern," *Mitteilungen Forschungsarbeit auf dem Gebeite Ingenieurwesens*, **24**, 1–51.

4

The Failure Theory for Isotropic Materials

All the results in this book are pivotally dependent upon the developments in this chapter. It contains the derivation (not postulation) of the failure theory for homogeneous and isotropic materials. The first two sections involve the construction of helpful steps that naturally lead to the derivation. The final two sections, after the three derivation sections, examine and reinforce the generality of the resultant theory of failure.

4.1 Theoretical and Testing Problems

The four standard failure criteria—Mises, Tresca, Coulomb–Mohr, and Drucker–Prager—are completely disqualified from serving as general failure criteria applicable to all homogeneous, isotropic materials. The first two are one-property forms and the second two are two property forms, none of which succeed in gaining generality, as was shown in Chapter 3.

It has gradually and tacitly become assumed that because of their failure to function as general failure criteria, the elusive general theory must be of a three or more property form. Most modern attempts at general theories have accordingly involved three or more properties.

What is more, with the three or more property forms, they are still usually "validated" by comparison with failure data for a particular materials type. But the associated testing is excruciatingly difficult, and the resultant data are commensurately scattered. It is very difficult to establish any degree of generality by this method.

All of this constitutes the nature and scale of the failure problem. The modern consensus of the likelihood for a unified, general failure theory seems to be that all the uncertainties and barriers place it as being somewhere between extremely difficult and impossible.

The assessment here is that the empirical three or four or more properties approach is just that—an empirical approach that never can and never will lead to generality. When viewed in this manner, three or more properties are not really properties at all; they are usually just fitting parameters to be adjusted to minimize the error with data over a specified range.

If there is to be any success in developing a general theory of failure there must be a clear-cut differentiation between what could be called a "physical properties" approach and what is merely a "parameters fit" to highly variable, limited sets of data.

4.2 Properties or Parameters

Part of the longstanding ambiguity and confusion with failure criteria/failure theory is due to the following. It is quite easy with three or more parameters to fit a limited range of failure data for a particular material or a particular class of materials. Often when that is done it becomes easily misinterpreted as being a proof of generality when it actually is merely an empirical data-fitting operation.

It takes a completely different method of verification to establish generality and proof over a comprehensive range. First, there must be success in agreeing with high-quality data for all the major materials classes, not just selected ones. Secondly, any purported general form must admit physically reasonable and proper behavior at the ductile and brittle ends of the scale. In other words, a comprehensive failure theory must admit the proper limiting case behaviors at the perfectly ductile limit and at the perfectly brittle limit. Furthermore, these limits must be defined with great care. Thirdly, it must be recognized that the basic failure theory itself must have a solid and rational foundation. Any form based purely upon conjecture or intuition would be likely to fail as being a legitimate failure theory.

Whichever of these two approaches is followed determines whether the embedded calibrating coefficients are actually properties or just parameters. Interest here is only with the properties situation. The derivation to be developed is based on the premise that the failure theory is in effect the constitutive relation of failure. Just as the constitutive stress–strain relations for elasticity are part of the materials constitution, so too

are its failure characteristics. This will be shown to lead to a particular and specific approach for the failure-theory problem.

It will be shown further that just as linear elasticity theory has two independent constitutive properties, so also does the coordinated failure theory have just two independent constitutive properties, when it is viewed as prescribing the range-limiting termination of elastic behavior. The expectation is that if a general theory of failure is to be found, it will necessarily have but two independent properties, not three or four or more.

The objective is to develop a two-property theory of failure for isotropic and homogeneous materials. If it happens that this cannot be achieved to satisfy the conditions stated above, then it will mean that a general theory cannot be found and probably does not exist. In that case, recourse can only be made to the three or more parameters empirical approach involving fitting data on a case-by-case basis. This specific point of departure marks the crucial junction between following one or other of these two distinctly different pathways.

4.3 The Organizing Principle

For a general approach to failure theory to be successful, it could be immeasurably helpful if it originated from some unifying concept or some special type of physical identification. In searching for an organizing principle for failure criteria, there is an overwhelming example that at least supplies the motivation and inspiration for trying to do so. This is the Periodic Table of the elements. Its organizing principle is that of the sequence of atomic numbers that then admits further classification when arranged in table form, with the rows representing the numbers of shells containing electrons, and the columns designating the numbers of electrons in the outermost one or, in some cases, two orbits.

Is there such a unifying and organizing basis for materials failure? In pursuing this possibility, attention will be focused upon isotropic materials. In scanning across the various materials types, there is an apparent, even obvious differentiation, but still with a connecting thread between them. This difference is that some types of materials have about the same strength in simple tension as in compression, while some others

have far less strength in tension than in compression. Still other groups fall between these two groups.

Since the times of antiquity it has always been well known and well utilized that gold falls into the first group. It also was learned how to live with the restrictions imposed on the second group for all the common minerals used in ancient construction. However, in the era of failure criteria activity, the realization of the tensile-versus-compressive strength differences never progressed beyond that point. There only was an awareness of the negative consequences of the possible disparity, and no recognition of a possible deeper meaning.

The present approach to failure characterization is based upon the physical observation/hypothesis that the general isotropic failure behavior is completely organized by and determined by the spectrum of T/C values, where T and C are the uniaxial tensile and compressive strengths. The spectrum of T/C values captures the entire spread from brittle behavior to ductile behavior over the range

$$0 \leq \frac{T}{C} \leq 1$$

This brittle-to-ductile change with the variation of the T/Cs is an integral and important part of the organizing structure. In fact, it provides the portal to all further developments to be presented here.

The T/C ratio at a particular value will be referred to as the "materials type." Thus there is a continuous spectrum of materials types, not just a sequence of discrete groups. The traditional materials groups, polymers, ceramics, and so on, have rather specific bands of T/Cs, although there can be, and are, overlaps.

The individual strength properties T and C have very different roles and functions to fulfill in synthesizing the related failure theory. Both T and C are involved (but unspecified) at the ductile limit of $T/C = 1$. In contrast, only C is involved (but unspecified) at the brittle limit, $T/C = 0$, because $T \rightarrow 0$. Accordingly, it is advantageous to nondimensionalize stress with respect to C (using T would not be possible) and then ideally have the ensuing failure criterion be expressed entirely in terms of the T/C spectrum. Following this course, take as the primary stress variable

$$\hat{\sigma}_{ij} = \frac{\sigma_{ij}}{C}$$

This organization of failure theory based upon the T/C spectrum does not automatically determine the failure criterion. Nevertheless, it does provisionally but strongly indicate that only two properties should suffice for the specification. More specifically, with T/C being the material classification to be followed, with no exceptions or ambiguity, then it must be a two-property failure theory involving only T and C. It cannot involve three or more properties (parameters). As speculated in Section 3.6, this two-property failure theory may in fact be quadratic in nature.

4.4 The Constitutive Equations of Failure, Part A: Polynomial-Invariants Criterion

The constitutive-equation approach will be used here to define failure as prescribing the effective limit of linear elastic behavior for all isotropic materials having a linear elastic range. The derivation follows from that of Christensen [4.1].

To get started, the first three invariants of σ_{ij} follow from the eigenvalue problem of the stress tensor, and are given by

$$\begin{aligned} I_1 &= \sigma_{ii} \\ I_2 &= \sigma_{ij}\sigma_{ij} \\ I_3 &= |\sigma_{ij}| \end{aligned} \tag{4.1}$$

Sometimes these are stated with the coefficients that appear in the eigenvalue problem, but as such are of no consequence here.

It is convenient to replace I_2 by J_2, with no loss in generality, where J_2 is expressed in terms of the deviatoric stress tensor, thus:

$$J_2 = s_{ij}s_{ij} \tag{4.2}$$

where

$$s_{ij} = \sigma_{ij} - \frac{\delta_{ij}}{3}\sigma_{kk} \tag{4.3}$$

For the purposes of representing various physical quantities, as yet unspecified, take ϕ as a polynomial expansion in the invariants, giving

$$\phi = \alpha I_1 + \beta I_1^2 + \gamma J_2 + - - - \tag{4.4}$$

The next three terms are of cubic order, and so on.

First consider the case where ϕ represents the elastic energy, U, for a linear, isotropic elastic material. Then

$$U = \alpha I_1 + \beta I_1^2 + \gamma J_2 + - - - \tag{4.5}$$

For a stress-free initial state and for positive definite energy the α term in (4.5) must be taken to vanish, leaving

$$U = \beta \sigma_{ii}^2 + \gamma s_{ij} s_{ij} \tag{4.6}$$

where the polynomial expansion is and must be truncated at the second-degree terms. The proper justification for the truncation involves normalizing the stresses by the Young's modulus, E, so that these terms are of the order of the strains and then the truncation represents the neglect of higher-order terms.

Evaluating β and γ to give the proper linear elastic form then leaves (4.6) as

$$U = \frac{1}{2E}\left[\frac{(1-2\nu)}{3}\sigma_{ii}^2 + (1+\nu)\, s_{ij} s_{ij}\right] \tag{4.7}$$

involving the two elastic properties E and ν. One can see at a glance from (4.7) that the necessary limits on ν are $-1 \leq \nu \leq 1/2$, otherwise the energy would not be positive definite. From the related variational theorem of linear elasticity, Sokolnikoff [4.2], there follows

$$\varepsilon_{ij} = \frac{\partial U}{\partial \sigma_{ij}} \tag{4.8}$$

which then gives the usual stress–strain relations when (4.7) is inserted into (4.8).

After deriving the linear elastic energy–stress–strain constitutive relations, attention is now turned to the constitutive relation for the failure of the linear elastic material. The failure-type constitutive relation would be expected to be related to but formally independent of the elastic energy–stress–strain forms. Note that the two terms—constitutive equations and constitutive relations—will be used interchangeably from here on.

The obvious possible criterion for failure would be that of a limiting value on the elastic energy U given by (4.7). This most direct approach has been exhaustively studied over the span of history. It is to no avail, as a limit on elastic energy could not possibly provide the failure constitutive relation, because isotropic materials do not fail under hydrostatic pressure conditions, at least not within the usual range of pressures for engineering materials.

Having disposed of that misguided notion, a more reasoned approach and derivation of the constitutive equation of failure will now be undertaken. The same formalism as is used to find the linear elastic constitutive relation for energy, (4.7), will be used to develop independently the constitutive relation for failure, viewing it as the specification for the range of validity of the energy form (4.7).

Moving along the direction just specified, take the constitutive relation of failure, which is to be coordinated with that for energy, as specified by

$$\phi \leq 1 \tag{4.9}$$

where this ϕ, now representing failure, will be given a parallel but separate polynomial expansion as that just used for energy. In performing this polynomial expansion in terms of the invariants of the stress tensor, the expansion must be truncated as the same level as that for the elastic energy, in order to coordinate with it. Thus the expansion (4.4) will be used in (4.9), properly truncated at second-degree terms. One more restriction must be imposed. For the materials of interest that do not fail in a state of hydrostatic compression, it is necessary to take $\beta = 0$ in the expansion (4.4). All of this then produces the basic constitutive relation for failure as

$$\alpha I_1 + \gamma J_2 \leq 1 \tag{4.10}$$

Following the guidelines provided in the preceding section on the organizing principle, the stress will be taken in the non-dimensionalized form of

$$\hat{\sigma}_{ij} = \frac{\sigma_{ij}}{C} \tag{4.11}$$

Then (4.1), (4.2), and (4.11) combined with (4.10) gives it as

$$\alpha\hat{\sigma}_{ii} + \gamma\hat{s}_{ij}\hat{s}_{ij} \leq 1 \tag{4.12}$$

where the α and γ in (4.12) are necessarily different from the like symbols in (4.10). Evaluating α and γ in (4.12) to give failure in uniaxial tension and compression as specified by T and C, gives (4.12) as

$$\left(\frac{C}{T} - 1\right)\hat{\sigma}_{ii} + \frac{3}{2}\left(\frac{C}{T}\right)\hat{s}_{ij}\hat{s}_{ij} \leq 1 \tag{4.13}$$

Finally, (4.13) can also be written in the form

$$\left(1 - \frac{T}{C}\right)\hat{\sigma}_{ii} + \frac{3}{2}\hat{s}_{ij}\hat{s}_{ij} \leq \frac{T}{C} \tag{4.14}$$

and where, as always,

$$0 \leq \frac{T}{C} \leq 1 \tag{4.15}$$

The highly compact form (4.14) is the main result of this derivation for failure theory. With the stress non-dimensionalized by the uniaxial compressive strength C, the only materials characteristic that enters the constitutive failure form is the ratio T/C. This ratio represents the materials type over the complete spectrum (4.15) of all materials types. This result and interpretation is the consequence of requiring that the failure constitutive relation represent the limiting range of the linear elastic behavior. In particular, the failure constitutive form (4.14) is the coordinating failure complement of the elastic energy constitutive form (4.7). Both are fully calibrated by uniaxial stress properties: namely, E and ν in (4.7) and their failure conjugates T and C in (4.14).

It is important to recognize that for materials that exhibit a significant plastic deformation after the elastic range is passed, the failure criterion (4.14) remains unchanged. The path of plastic deformation gives a more complex transition from the elastic state to failure than in the case of brittle materials, but it does not alter this special relationship between the two states of linear elasticity and failure. Whether there is abrupt

brittle failure or drawn out ductile failure the overall controlling main event is fundamentally the same thing; it is the cessation of the elastic behavior. The material can no longer sustain increasing internal loads of elastic deformation, and it fails, either suddenly or gradually. All this is explained further, and the means of defining failure stress (strength) from elastic–plastic data is taken up in Chapter 9.

Writing the constitutive relation for failure, (4.14), in terms of components gives

$$\left(1-\frac{T}{C}\right)(\hat{\sigma}_{11}+\hat{\sigma}_{22}+\hat{\sigma}_{33})+\frac{1}{2}\left[(\hat{\sigma}_{11}-\hat{\sigma}_{22})^2+(\hat{\sigma}_{22}-\hat{\sigma}_{33})^2+(\hat{\sigma}_{33}-\hat{\sigma}_{11})^2 + 6\left(\hat{\sigma}_{12}^2+\hat{\sigma}_{23}^2+\hat{\sigma}_{31}^2\right)\right]\leq\frac{T}{C} \tag{4.16}$$

In terms of principal stresses, (4.16) becomes

$$\left(1-\frac{T}{C}\right)\hat{\sigma}_{ii}+\frac{1}{2}\left[(\hat{\sigma}_1-\hat{\sigma}_2)^2+(\hat{\sigma}_2-\hat{\sigma}_3)^2+(\hat{\sigma}_3-\hat{\sigma}_1)^2\right]\leq\frac{T}{C} \tag{4.17}$$

For many purposes and problems it is necessary only to specify T/C as the materials type, and the failure stresses are simply scaled by C as in (4.11). This has great utility in applications.

In dimensional stress form the constitutive relation of failure (4.17) is given by

$$\left(\frac{1}{T}-\frac{1}{C}\right)\sigma_{ii}+\frac{1}{2TC}\left[(\sigma_1-\sigma_2)^2+(\sigma_2-\sigma_3)^2+(\sigma_3-\sigma_1)^2\right]\leq 1 \tag{4.18}$$

Henceforth these constitutive relations of failure—(4.14), (4.16), (4.17), and (4.18)—will be referred to as failure criteria, and all these forms will be designated as being of the polynomial-invariants type, since that is the manner of their derivation.

In principal stress space the geometric form of this failure criterion is that of a paraboloid as shown in Fig. 4.1.

The axis of symmetry of the paraboloid makes equal angles with the three principal stress axes. In the limiting case the paraboloid becomes a cylinder, as will be discussed in Section 4.6.

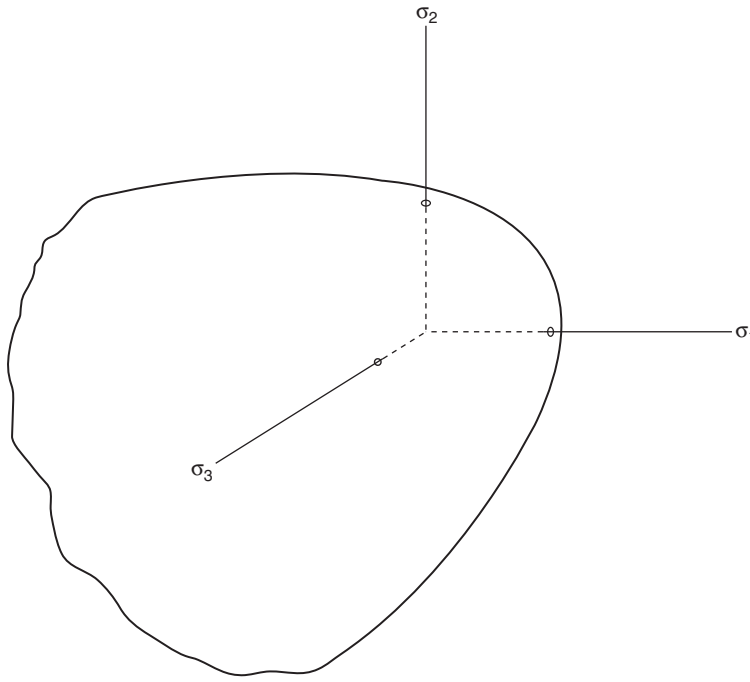


Fig. 4.1 *Polynomial-invariants paraboloid in principal stress space.*

Important though it is, this paraboloidal, polynomial-invariants criterion is not the complete and sole failure criterion that controls all possible failure occurrences. As explained in the next section, there is also a competitive fracture mode of failure that must be taken into account.

4.5 The Constitutive Equations of Failure, Part B: Fracture Criterion

The polynomial-invariants failure criterion just derived is that of an axisymmetric paraboloid in principal stress space. There is much evidence that this type of failure criterion is not sufficient for all cases over the full range of materials types. In fact, there is much anecdotal evidence that for brittle materials there exists three-fold symmetry rather than axial symmetry in the failure envelope form.

Another facet of possible problems is that the maximum normal stress criterion of Lamé and of Rankine seems to have some (but not total) validity in some stress states for brittle materials.

The most compelling evidence of a quite complex failure behavior is that running cracks (either slow or fast) in isotropic, homogeneous materials are usually observed to propagate as a Mode I crack in a direction normal to the maximum normal stress direction, as reported in many reliable sources such as Broberg [4.3]. It is entirely consistent with this evidence that there could exist a fracture mode of failure in homogeneous, elastic materials, and this related fracture criterion is somehow controlled (activated) by the maximum principal stress.

All of this appears to pose a major problem in the development of the constitutive equations of failure. If a maximum normal stress type fracture criterion is added to the polynomial-invariants criterion just derived, it would be expected to require additional parameters in order to specify it. This is precisely what has been the "dead-end" path followed by nearly all previous failure criteria developments. Typically, additional criteria would be added, bringing in new parameters. This approach then invariably degenerates into a parameter-fitting operation, as described and discussed earlier. In doing so, the contact with physical reality becomes ever more thin and tenuous.

There is only one consistent and legitimate condition for bringing in an additional fracture criterion, competitive with the polynomial-invariants criterion (4.14). This new fracture criterion must not add any new parameters to the process, nor can it violate any of the failure interpretations already established. Primarily it must be of effect in the brittle range of behavior, and not disturb the ductile side of the scale. The polynomial-invariants criterion has two failure properties as calibrating it. A new, competitive fracture criterion cannot add new parameters, bringing the total to three or more parameters.

One further and final requirement must be imposed upon any additional fracture criterion. The failure surface envelope in stress space generated by the polynomial-invariants criterion (4.14) is a continuous function of T/C. Correspondingly, the possible fracture criterion must not introduce any discontinuous form in the failure surface envelope as T/C is varied. This means that a fracture criterion cannot activate at some arbitrary value of T/C that then causes discontinuous behavior.

These conditions on an additional fracture criterion appear to be so impossibly restrictive as to disallow its occurrence. It will be found, however, that all the required and constrictive conditions of this section still permit one, and only one, special fracture criterion. This will

occur without upsetting the properties balance or the significance of the polynomial-invariants criterion or the other features of expected behavior already outlined.

The one fracture criterion that satisfies all of the forgoing conditions is as follows.
If

$$\frac{T}{C} \leq \frac{1}{2}$$

then

$$\begin{aligned} \sigma_1 &\leq T \quad \hat{\sigma}_1 \leq \frac{T}{C} \\ \sigma_2 &\leq T \text{ or } \hat{\sigma}_2 \leq \frac{T}{C} \\ \sigma_3 &\leq T \quad \hat{\sigma}_3 \leq \frac{T}{C} \end{aligned} \tag{4.19}$$

In the most compact form with σ_1 being the largest principal stress, then (4.19) can be written as

$$\hat{\sigma}_1 \leq \frac{T}{C} \quad \textit{for} \quad \frac{T}{C} \leq \frac{1}{2} \tag{4.20}$$

Whichever of the two failure criteria—the polynomial-invariants form (4.17) or the fracture mode form (4.19)—is the more limiting is then the controlling failure condition at that stress state. The geometry of the failure surfaces for these two failure modes will be briefly discussed next.

The fracture criterion (4.19) is that of three planes normal to the three coordinate axes in principal stress space. It is at the value of $T/C = 1/2$ that the three fracture-controlled planes are just tangent to the polynomial-invariants paraboloid. This geometric characteristic ensures that there will not be any physical discontinuities in behavior as T/C is varied across the region where the fracture criterion commences, $T/C = 1/2$. For $T/C < 1/2$ the three fracture planes cut sections out of the paraboloid. For these values of T/C very near to $1/2$, these sections or slices cut from the paraboloid are very small, but they increase in size

as T/C diminishes. All these effects will be illustrated and amply discussed in Chapter 5, and the experimental verification will be taken up in Chapter 6.

The final, general statement of failure criteria for all homogeneous and isotropic materials thus requires the imposition of two separate but coordinated and competitive criteria. These are the forgoing polynomial-invariants criterion and the fracture criterion of this section. It is seen from the derivation of these constitutive relations of failure that the polynomial-invariants result is necessarily primal and provides the "backbone" for the overall method. It reveals the special relationship between the elastic energy constitutive relation and the polynomial-invariants constitutive relation (of failure), as one without the other is incomplete.

Nevertheless, the fracture criterion is not just a minor adjunct to the polynomial-invariants criterion. In addition to providing more restrictive results in some cases, the fracture criterion is absolutely necessary for establishing the generality of the entire failure methodology. Furthermore, as elaborated below, it will provide the key to enlightenment on all matters related to ductile-versus-brittle failure behaviors.

This completes the derivation of the constitutive relations of failure for homogeneous and isotropic materials. Major parts of this book are spent discussing the implications and applications of these results. For example, it is seen that the fracture criterion is active for materials with $T/C < 1/2$, but it is not relevant to materials with $T/C > 1/2$. This will be found to have profound implications for the ductile and brittle behaviors of the failure modes. However, the situation is far more subtle and complex than simply claiming that all materials with $T/C > 1/2$ are ductile while all those with $T/C < 1/2$ are brittle. Some aspects of the ductile/brittle problem will be discussed in Chapter 5, while the full ductile/brittle landscape is surveyed and assessed in Chapter 8.

Before closing this derivation chapter, the limiting cases of these failure criteria will be examined closely.

4.6 Ductile and Brittle Limits

Now consider the ductile and brittle limits that any general failure criterion should include. The ductile limit is easy and obvious. Most would agree that it must be the Mises criterion, while a few might prefer the

Tresca criterion. Between these two longstanding forms the ductile limit case is covered. However, the situation with the brittle limit is historically completely opaque.

There is not even agreement on what should constitute the brittle limit, nor if such a thing even exists. In the two-property polynomial-invariants method presented here for homogeneous and isotropic materials, the uniaxial tensile and compressive strengths, T and C, are found such that the ductile limit is defined by $T/C \to 1$, while the brittle limit is defined by $T/C \to 0$. Some would say that the brittle limit so specified is only a mathematical abstraction and of no practical use or interest. Such is not the case.

It is argued here that for any candidate failure criterion it is very important to examine its behavior in the brittle limit as $T/C \to 0$ in order to determine whether reasonable and consistent results emerge, or if physically unacceptable or pathological results are obtained. In the latter case the supposed failure criterion should be discarded, because with aberrant behavior at the limit it would be certain that the form could not have generality and would be of little or no value.

For example, consider the two-property Coulomb–Mohr failure criterion. In the brittle limit, $T/C \to 0$, the Coulomb–Mohr form allows all principal stresses that are negative (or zero) in value. This then implies that such a material has a uniaxial compressive strength that is of indeterminate size, and even unlimited in magnitude. In effect, no property is involved at the brittle limit, yet the material can support load. This clearly is an unphysical and unacceptable behavior, as the Coulomb–Mohr failure criterion is degenerate at the brittle limit. This is one of the simplest ways of revealing the general unacceptability of the Coulomb–Mohr criterion. Other examples of proposed but actually inadmissible failure forms will be presented later.

As a counter-example, consider the failure criterion developed in the preceding three sections. First, by way of a little background, the Mises criterion has only one property. For a general two-property failure theory that has the Mises criterion at its ductile limit, one might reasonably expect that its brittle limit would also turn out to be a one-property form. As shown below, exactly that does occur in this theory.

The governing equations for the present two-property isotropic material failure criterion are given by (4.16)–(4.19). In the ductile and brittle limits and in terms of principal stresses, these become

Ductile limit (Mises), $T = C, \quad S = T/\sqrt{3}$

$$\frac{1}{2}\left[(\sigma_1 - \sigma_2)^2 + (\sigma_2 - \sigma_3)^2 + (\sigma_3 - \sigma_1)^2\right] \leq T^2 \tag{4.21}$$

Brittle limit, $C \neq 0$, $T = S = 0$

$$C\,(\sigma_1 + \sigma_2 + \sigma_3) + \frac{1}{2}\left[(\sigma_1 - \sigma_2)^2 + (\sigma_2 - \sigma_3)^2 + (\sigma_3 - \sigma_1)^2\right] \leq 0 \tag{4.22}$$

and

$$\begin{aligned} \sigma_1 &\leq 0 \\ \sigma_2 &\leq 0 \\ \sigma_3 &\leq 0 \end{aligned} \tag{4.23}$$

where S is the shear stress at failure. The ductile limit is seen to be that of the Mises criterion, and it takes the form of a cylindrical surface in principal stress space.

The brittle limit is shown in Fig. 4.2 in three-dimensional graphics form. The coordinate axes are given in the negative stress directions.

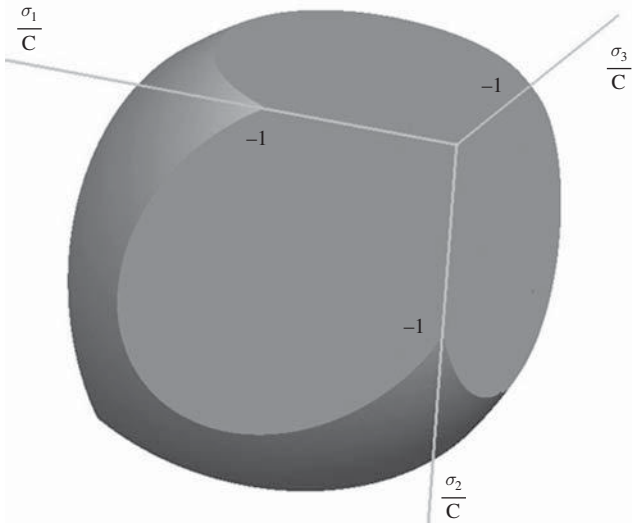


Fig. 4.2 *The brittle limit in principal stress space.*

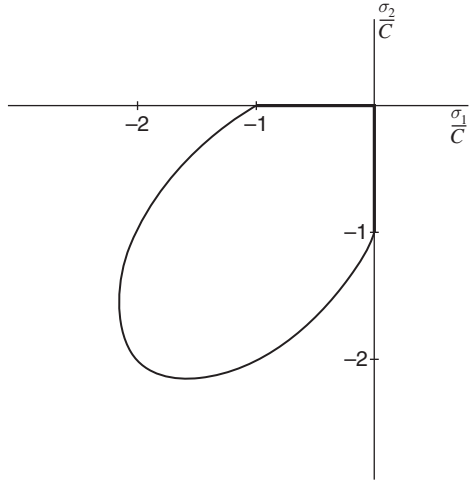


Fig. 4.3 *The brittle limit for biaxial stresses.*

The general form of the failure surface is that of a triangular pyramid that transitions into a paraboloid with corners between the two regions, and it can be seen that over much of the region there is three-fold symmetry, as expected.

Taking $\sigma_3 = 0$ gives the two-dimensional biaxial stress failure envelope at the brittle limit, as in Fig. 4.3. It is interesting to note that at the brittle limit in equal biaxial compression the failure level is exactly twice that in uniaxial compression. At the ductile limit they are the same. This is the first "hint" of the stabilizing effect of hydrostatic pressure in damaged materials. Another example will be presented below.

It does not follow that the failure modes are necessarily brittle in these two-dimensional and three-dimensional examples. The large negative mean normal stresses have an influence on that situation. The significance of these results is that the brittle limit does not allow any tensile stresses at failure, in addition to the other primary features.

The failure criteria at the brittle limit show that the shear strength, S, must vanish. However, if a state of pressure, p, is applied first, then some shear strength would become possible, as a superimposed stress state. This is found to be true, with the shear strength dependence on the pre-applied pressure as shown in Fig. 4.4.

These examples and results demonstrate that a well-founded three-dimensional failure theory can and must evoke consistent behavior at the

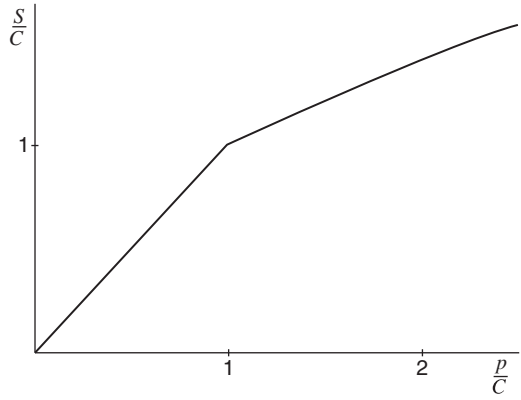


Fig. 4.4 *Effect of pressure on shear strength at the brittle limit.*

brittle limit. The explicit physical form of the brittle limit is also of special interest in its own right.

4.7 Problem Sets Purpose

Problem sets are included in this chapter and in all subsequent chapters under the heading "Problem Areas for Study". These problems are not the usual kind of "plug-in" exercises; nor do they refer to particular applications, whether in the past or not yet addressed. Rather, the problem areas relate to the basic formulation of failure theory and its assimilation. There are major openings for further work and for entirely new work on the many different aspects of failure criteria and their means of application.

The total collection of all the problem areas in all chapters is far from being exhaustive or even comprehensive. It is more like a sampling of possible problem areas in moving forward. No attempt is made to enter into the details of any of these areas, but they should convey the kernels of the problems. Some problems may be rather easy, but have simply not yet come to the surface. Others may be extremely difficult or even defy solution.

Before getting into the first problem set, a very brief account will be presented of the formative developments in the field of failure theory and failure criteria. The time line of main developments is particularly

relevant. This will provide the obvious prelude to further activities in general, and to the problem areas' opportunities in particular.

Although failure criteria have been around since the beginning of mechanics of materials, it has been only within the past many years that it has become more than just a casual speculation. In this regard there are interesting parallels and contrasts between the respective developments of failure criteria and fracture mechanics. Griffith launched fracture mechanics in the 1920s; and around the same time, Mohr took Coulomb's original concept of failure and presented it in an appealing and apparently usable form: the Mohr–Coulomb or Coulomb–Mohr criterion. So both efforts—failure criteria and fracture mechanics—effectively commenced almost simultaneously.

Griffith's approach to fracture languished for several decades before being revitalized and activated. In contrast, almost immediately there was high promise and even excitement for the Coulomb–Mohr theory. So much for short-term promise! The Coulomb–Mohr criterion gradually faded into the background as being only a historical artifact. It completely failed to demonstrate relevance for a wide range of isotropic materials. On the other hand, fracture mechanics eventually matured into the complete and active field that it is now. There matters stood until comparatively recent times, and it is only in the past few decades that failure criteria have returned to prominence and serious activity.

So in reality, failure criteria is a young technical enterprise. It is much younger in a developmental sense than the now mature field of fracture mechanics. When one says "fracture mechanics" there is absolutely no question as to the exact and explicit governing discipline. Chapter 10 outlines fracture mechanics and the distinction between it and failure criteria. When one says "failure criteria" a broad range of possible images quickly comes to mind. There are in circulation a great many empirical formulas of uncertain validity, but they all still fall under a broad subject heading of being some form of failure criteria.

However, the terminology "failure theory" and "failure criteria" is used here in a much more restrictive and specific sense. As should be apparent from the contents of this chapter, failure criteria herein refers to forms that have a significant theoretical basis for their existence and a physical compatibility with quality-testing data, when such data exist. Ideal though this may seem, making lasting judgments and commitments consistent with this objective, upon which everyone can agree, are never easy

nor quick. Caution and care are needed, and the problems stated here are intended to aid this process.

Problem Areas for Study

1. Are materials failures essentially and predominately untreatable anomalies, or are they amenable to general characterization through failure criteria as part of the constitutive specification? The latter direction is the course followed here. Are there any alternative or "in-between" approaches? If so, how should failure criteria be viewed as fitting into the overall formulation for materials performance?
2. The failure theories in this book involve surfaces in stress space that include intersections (corners). These result from the competition between failure modes. To what extent do these corners reflect reality, or are they just acceptable idealizations? Are acute-angle intersections less probable than wider-range angles?
3. The one-dimensional elastic stress–strain relation and the limiting failure stress(es) are obvious, and can be written down by inspection. However, the one-dimensional case can be formalized by the same method as was used in the three-dimensional derivation of Section 4.4, and then the results from this method are not so obvious. The differences between the steps of the derivation in the one-dimensional case and the three-dimensional case are significant, not trivial, but the end results are completely compatible. Probe this one-dimensional and three-dimensional comparison, and draw your own conclusions.
4. This chapter provides a two-property failure criterion for isotropic materials. The only other general purpose two-property forms are the Coulomb–Mohr criterion and the Drucker–Prager criterion. All other general forms that have been proposed involve three or more parameters. A contrary point of view would be that a two-property form is not possible. Is there a proof of this, or can a fundamental contradiction be constructed?
5. Relative to the present theory of failure, the brittle limit has been extracted in mathematical form in the Section 4.6. In principal stress space the failure form is shown in Fig. 4.2. Can a physical model involving damage patterns be reasoned to represent this case? If not in three dimensions, can it be achieved in two dimensions?

6. The basic question continues as to whether stress or strain is the appropriate variable for failure criteria. A rationale for stress has been presented here. What are the other supporting arguments for stress, and what are the arguments for using strain. Can either be reasoned to be truly fundamental?
7. How can the type of failure criteria derived and used here be modified to accommodate thermal stresses and residual stresses? Are there any thermodynamical insights to be brought to bear on these problems?

References

[4.1] Christensen, R. M. (2007). "A Comprehensive Theory of Yielding and Failure for Isotropic Materials," *J. Engr. Mater. and Technol.*, **129**, 173–81.

[4.2] Sokolnikoff, I. S. (1983). *Mathematical Theory of Elasticity*, 2nd ed., Krieger, Malabar, FL.

[4.3] Broberg, K. B. (1999). *Cracks and Fracture*, Academic Press, New York.

5
Isotropic Materials Failure Behavior

The present chapter and the subsequent two chapters are immediate and necessary continuations of the preceding chapter on the derivation of failure theory and failure criteria for isotropic materials. No matter how rigorous a derivation may be, it is of little value if it is not readily assimilated and readily interpretable and if it does not offer direct facility in evaluation and in applications. Several different two-dimensional and three-dimensional stress states with associated failure envelopes will be presented and discussed here to broaden the understanding. Other conditions and decision junctures will also be introduced: for example, failure types, such as ductile or brittle failure.

The failure criteria for isotropic materials are given by eqs. (4.18) and (4.19) in Chapter 4. As specialized to principal stresses they are

$$\left(\frac{1}{T}-\frac{1}{C}\right)(\sigma_1+\sigma_2+\sigma_3) + \frac{1}{2TC}\left[(\sigma_1-\sigma_2)^2+(\sigma_2-\sigma_3)^2+(\sigma_3-\sigma_1)^2\right] \leq 1 \tag{5.1}$$

and if

$$\frac{T}{C} \leq \frac{1}{2}$$

the following fracture criteria also apply:

$$\begin{aligned} \sigma_1 &\leq T \\ \sigma_2 &\leq T \\ \sigma_3 &\leq T \end{aligned} \tag{5.2}$$

The criterion (5.1) is referred to as the polynomial-invariants form, after its method of derivation.

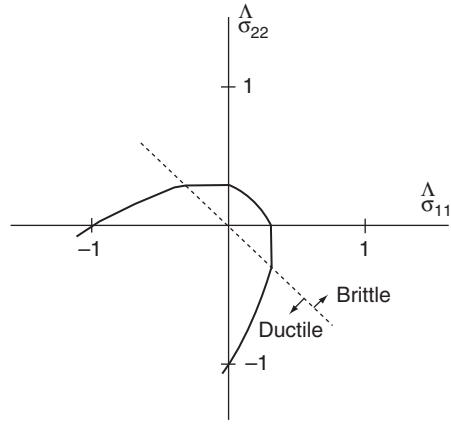


Fig. 5.1 *Biaxial stress failure, T/C = 1/3.*

5.1 Failure Behavior in Two Dimensions

As mentioned earlier, the fracture criteria (5.2) intersect the paraboloid (5.1) and produce three flattened surfaces on it. This is most easily illustrated in two-dimensional stress space σ_{11} and σ_{22} with the other stresses vanishing, Fig. 5.1, which is for $T/C = 1/3$. The stresses are non-dimensionalized by dividing by C.

Referring to the biaxial stress state,the $T/C = 1$ case would be the usual Mises ellipse. At $T/C = 1/2$ the fracture criterion (5.2) is just tangent to the shifted and slightly smaller ellipse (5.1) at the intercepts with the axes. The $T/C = 1/3$ case, Fig. 5.1, is typical of cast iron and some ceramics, showing a very pronounced fracture cutoff effect. The limiting case $T/C \rightarrow 0$ reveals that applied stress can be sustained only if the mean normal stress is negative and no component of normal stress is positive. The intuitive division into ductile and brittle regions shown in Fig. 5.1 will be found later to actually have a physical basis.

5.2 Failure in Principal Stress Space

The complete failure surface in three-dimensional principal stress space is composed of two parts: the paraboloid (as determined by polynomial invariants), and a competing fracture mode of failure that produces planar "fracture cutoffs" from the paraboloid. Mainly the geometric aspects of the failure surfaces will be shown and described here. The physical and

mathematical bases of these are the subjects of nearly the entire book, and will not be immediately involved here.

The complete failure criteria are given by eqs. (5.1) and (5.2). These failure criteria are calibrated by two failure properties, the uniaxial tensile and compressive strengths T and C. The paraboloidal part of the overall failure criterion, from the method of polynomial invariants, has a symmetry axis that makes equal angles with the three principal stress axes. The fracture criterion planes are normal to the three principal stress axes.

In the following graphics, stress is non-dimensionalized by the uniaxial compressive strength, C. The examples differ only through the variation of the tensile/compressive strength ratio, T/C. The different T/C values correspond to seven examples from widely different classes of materials: namely, ductile metals, ductile polymers, brittle polymers, cast irons, ceramics, glasses, and geological materials.

All of these cases are at the same scale, thus allowing direct comparisons between them. In this sequence of computer graphics the scale and the view angle are selected to best observe the brittle end of the sequence rather than the ductile end, since the latter is quite well understood. The ductile cases will be further illustrated later. A grouping of small-sized images is shown first in Fig. 5.2, followed by larger, individual images in Fig. 5.3a–5.3g. In some cases the stress axes are shown only in the negative coordinate directions.

$$\text{Coordinate Axes: } \frac{\sigma_1}{C},\ \frac{\sigma_2}{C},\ \frac{\sigma_3}{C}$$

The two shadings refer to the two different and competing failure criteria. They do not imply a ductile-versus-brittle distinction, which will be brought in later.

The unusual scaling for the Mises criterion shown in Fig. 5.3a is necessary for the one-to-one comparison capability with the other cases. When displayed by itself the Mises criterion usually is taken in a long, slender cylindrical form, as shown in Fig. 5.4.

The Mises criterion at $T/C = 1$ represents one of the limiting cases of failure behavior, the ductile limit. Its cylindrical form is the limiting case of the general paraboloid. The other limiting case of failure is that of the brittle limit, at $T/C = 0$. This case is similar to the geological materials graphic, being only slightly different when taken to the limit where the apex of the paraboloid coincides with the coordinate origin. Also at this

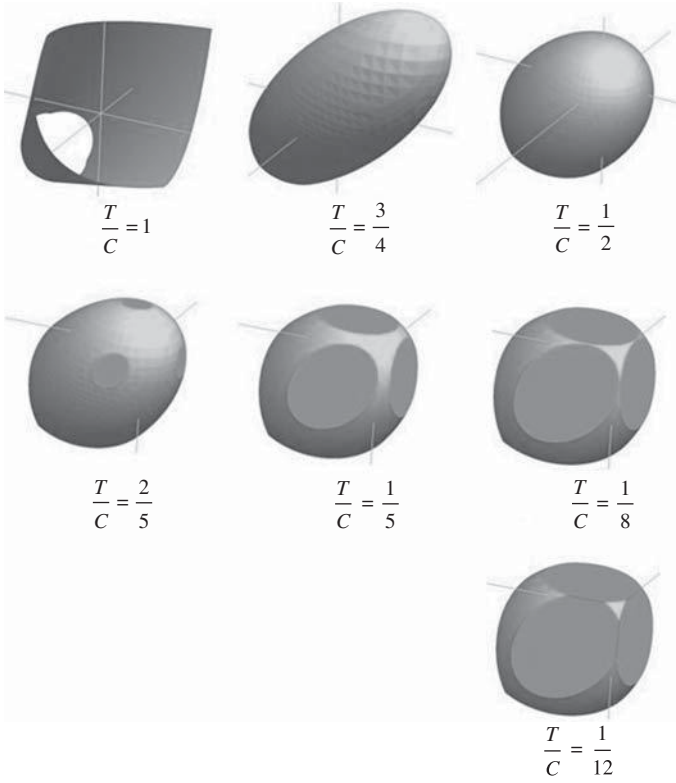
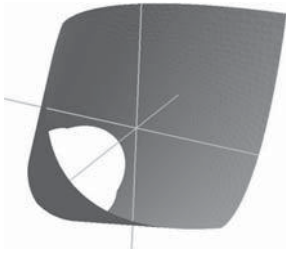


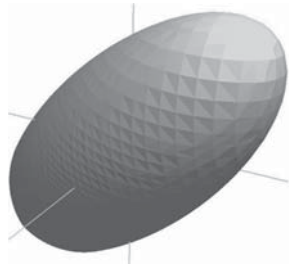
Fig. 5.2 *Spectrum of failure envelope types.*

limit, each of the three fracture planes contains the coordinate origin along with two of the axes, thereby requiring that no tensile component of three-dimensional stress exist anywhere in the domain. The brittle limit failure surface is shown in Fig. 5.5. This form has a triangular pyramid shape near the origin, then transitioning into the paraboloid with corners between the two, as shown.

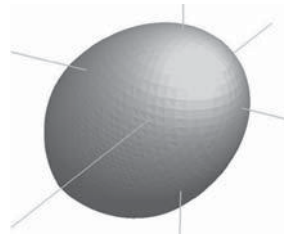
The two limiting cases of failure are remarkably different—indeed, fundamentally different. The most distinctive features of the ductile limit failure criterion are its independence of hydrostatic stress and its symmetry of a type in tensile and compressive stresses. In contrast, the most distinctive features of the brittle limit failure criterion are a strong dependence upon hydrostatic stress and the disallowance of any tensile components of stress, giving it a special type of asymmetry.



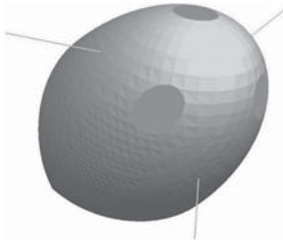
Ductile Metals (Mises Cylinder), $\frac{T}{C} = 1$



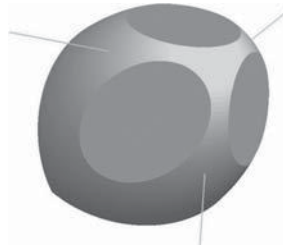
Ductile Polymers, $\frac{T}{C} = \frac{3}{4}$



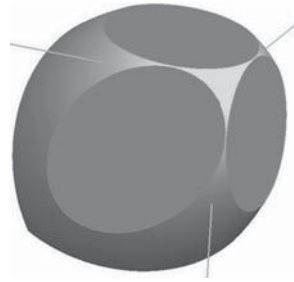
Brittle Polymers, $\frac{T}{C} = \frac{1}{2}$



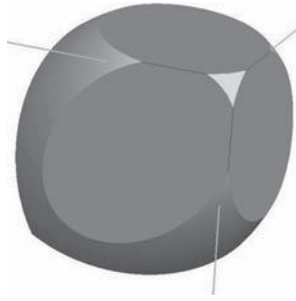
Cast Irons, $\frac{T}{C} = \frac{2}{5}$



Ceramics, $\frac{T}{C} = \frac{1}{5}$



Glasses, $\frac{T}{C} = \frac{1}{8}$



Geological Materials, $\frac{T}{C} = \frac{1}{12}$

Fig. 5.3a–g *Materials failure examples.*

At a different scale and view angle, the ductile polymers case at $T/C = 3/4$ appears as in Fig. 5.6. While not having the capability of the ductile limit, the form in Fig. 5.6 still reveals the very considerable tensile stress capabilities for this class of materials. Rather than examine other specific cases, some general geometric characteristics will now be noted.

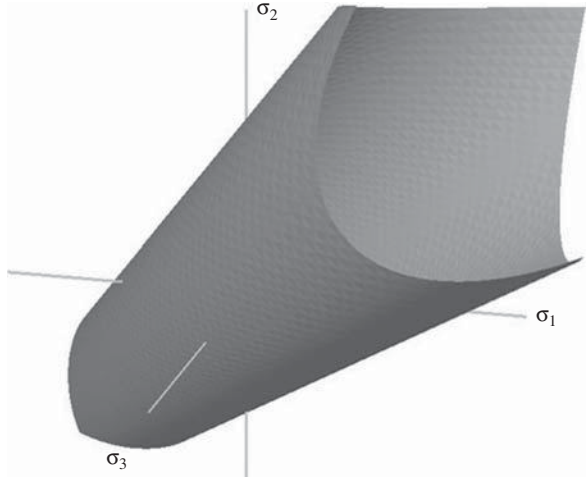


Fig. 5.4 *Mises cylinder.*

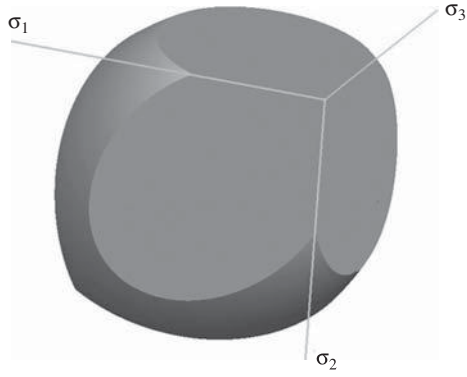


Fig. 5.5 *Brittle limit.*

In the examples and in general, the complete failure surface is axisymmetric for $T/C \geq 1/2$, but for $T/C < 1/2$ it has three-fold symmetry. It is at and below $T/C = 1/2$ that brittle behavior becomes much more prevalent with the onset of the fracture criterion. Even though the character of the failure surfaces changes at $T/C = 1/2$, there is no discontinuity of form for the failure surface at this value of T/C.

As T/C diminishes the apex of the paraboloid moves toward the coordinate origin. For $T/C < 1/2$ the intersection of the fracture planes with the paraboloidal surface generates elliptical forms. These elliptical

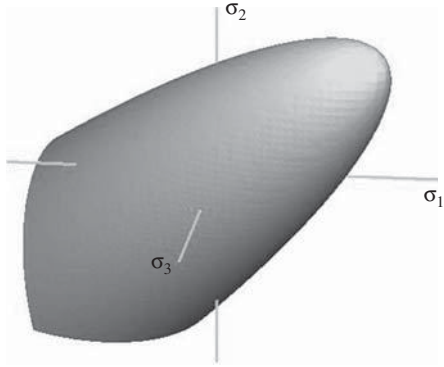


Fig. 5.6 *Ductile polymers.*

intersections become larger as T/C diminishes from the value of 1/2. Going in the other direction the three elliptical intersections reduce to points at $T/C = 1/2$, because the fracture planes become tangent to the paraboloid.

For a sufficiently small value of T/C the adjacent elliptical intersections become tangent with each other. For even smaller values of T/C the elliptical intersections become truncated and contiguous, as shown in Fig. 5.3g, the $T/C = 1/12$ graphic. The point of tangency occurs at

$$\frac{T}{C} = \frac{5 - 2\sqrt{3}}{13} = \frac{1}{8.46}$$

These examples show the intricate and non-intuitive failure surface changes that occur in going across the full spectrum of isotropic materials types. They also show how extremely limited the allowable tensile stress regions become relative to the allowable compressive stress regions as T/C decreases.

Although the fracture cutoffs are nominally brittle, it does not follow that the remaining paraboloidal part of the failure surface is entirely ductile. For any particular value of T/C there is a mean normal stress controlled plane which is normal to the symmetry axis of the paraboloid and which divides the domain into ductile and brittle parts. This ductile/brittle division plane cuts across the intersection ellipses, where they exist.

Despite the exceptionally wide range of failure surface forms for isotropic materials, all are determined by and calibrated by only two failure properties: the uniaxial tensile strength and the uniaxial compressive strength.

5.3 Ductile-versus-Brittle Failure

The distinction between ductile failure and brittle failure is one of the cornerstones of materials science, in addition to being of utmost importance for materials applications. Ductile–brittle transitions as influenced by some control variable, such as temperature, have been very extensively examined. Basic theoretical studies of the transition effect have usually (but not always) been posed in the context of the activation and flow of atomic scale dislocations in face-centered cubic and body-centered cubic metals, or in more idealized forms. Despite the large amount of research already carried out on the ductile–brittle transition, very little of it has been successfully reduced to standard practice and a uniformity of understanding for the full spectrum of materials types. The specification of ductile-versus-brittle failure criteria for all homogeneous materials has continued to be indistinct and elusive.

Some perspective on the problem can be gained as follows. The same as with temperature, pressure can be used to control ductile–brittle effects. Sufficiently large superimposed pressure can convert what is normally thought of as a brittle material into a ductile material. Bridgeman was awarded the Nobel Prize at least partially for illuminating the problem nearly a hundred years ago. However, positive or negative pressures are but a part of the full stress tensor, and because of this they are more complex than scalar temperature variation. This complexity has been something of a barrier to the understanding of ductile–brittle failure effects.

Proceeding further with this ductile–brittle reconnaissance, all stress states can be decomposed into a dilatational part (as with hydrostatic stress) and a distortional part (as with shear stress). Consider the three following stress states: uniaxial tension, eqi-biaxial tension, and eqi-triaxial tension. The decomposition of these stress states into dilatational, and distortional parts are as shown in Table 5.1.

Each of the two distortional (shear) stress states can be further decomposed into two simple shear stress states, as with

$$\sigma_{11} = 1/3, \sigma_{22} = -1/3, \sigma_{33} = 0, \text{ and } \sigma_{11} = 1/3, \sigma_{22} = 0, \sigma_{33} = -1/3$$

Table 5.1 Uniaxial tension, eqi-biaxial tension, and eqi-triaxial tension

		Dilatational	Distortional (shear)	Total
Uniaxial tension				
	σ_{11}	1/3	2/3	1
	σ_{22}	1/3	–1/3	0
	σ_{33}	1/3	–1/3	0
Eqi-biaxial tension				
	σ_{11}	2/3	–2/3	0
	σ_{22}	2/3	1/3	1
	σ_{33}	2/3	1/3	1
Eqi-triaxial tension				
	$\sigma_{11} = \sigma_{22} = \sigma_{33} = \tilde{T}$, no shear			

For further investigation, consider only those materials that are normally thought of as being very ductile, such as aluminum, copper, and so on. Such isotropic materials have the uniaxial tensile and compressive yield stresses as being identical, $T = C$, to within experimental discrimination. Also, the same failure values occur for eqi-biaxial tension, for these materials. Other materials types have $T/C < 1$, or much less in some cases. For convenience, refer to the distortional part of the stress tensor as shear. Designate by S a measure of the shear part of the stress tensor (the Mises stress for example), and by D designate a measure of the dilatational part, both at yield/failure.

Normally it is only the shear part that controls yield/failure in these $T = C$ materials, but it is both S and D that relate to the ductile-versus-brittle failure mode type. For present purposes, take a possible ductility index for these three cases as specified by S/D. In the uniaxial tension case, normalize this S/D ratio to the value of 1. It then follows that the S/D ratio for the eqi-biaxial tension case will have the value of 1/2, because it has the same value for S as in uniaxial tension (the sign change shown in the table is irrelevant) but its value for D is twice as large as that in uniaxial tension. Finally, for the eqi-triaxial tension case there is no shear, $S = 0$, so its ductility index is $S/D = 0$.

In this manner and for these three cases there is a ductility scale that goes from 0 to 1, with the value 0 being that of no ductility (major brittleness) and that of 1 being major ductility. Collecting these three cases for materials with $T = C$ then gives

Uniaxial Tension	$S/D = 1$, very ductile
Eqi-Biaxial Tension	$S/D = 1/2$
Eqi-Triaxial Tension	$S/D = 0$, very brittle

The triaxial case falls into the brittle category, since there are no shear stresses to activate dislocation flow in these $T = C$ materials. That an eqi-triaxial tension test is exceedingly difficult to perform is not of concern here, as only the type of failure is being considered. Also, for this simple exercise, the fact that the idealized Mises criterion would say there is no failure under triaxial tension is of no relevance.

The present reasoning and discussion is for background purposes only, and does not represent a specific technical approach. Nevertheless, it does quite clearly indicate that there is a primary linkage between (i) ductile-versus-brittle failure behavior and (ii) the type of stress state which is imposed. Any comprehensive failure theory must account for these effects. They can be just as important as the effects due to temperature variation. They comprise the litmus tests (criteria) indicating and differentiating ductile yield behavior from brittle failure behavior.

A general criterion for the ductile–brittle transition is developed in Chapter 8. The theory given there interestingly determines that the state of eqi-biaxial tension for materials with $T = C$ does in fact lie on the borderline transition between ductile yield and brittle failure. This same effect is again shown in Fig. 5.7, as a preview of Chapter 8. More importantly, the theory covers all isotropic materials types having T $\leq$ C, and all stress states.

The failure criteria (5.1) and (5.2) have been examined in considerable detail and developed into a criterion for ductile-versus-brittle behavior.

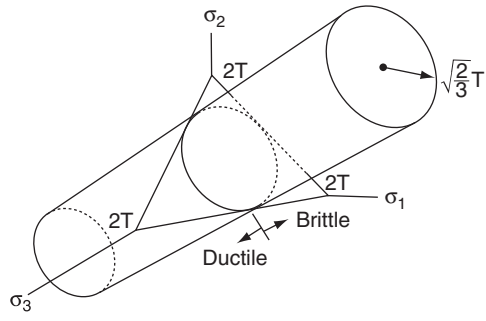


Fig. 5.7 *Mises cylinder, principal stress space.*

Thus the modified paraboloid in stress space is subdivided into regions of brittle failure versus regions of ductile failure. The brittle regions are not just the fracture planes produced by (5.2), but also include portions of the paraboloid (5.1). For given values of T and C the specific stress state which is imposed determines whether the failure will be of ductile or of brittle nature. The resultant stress state division into ductile and brittle regions would not actually be expected to have a sharp dividing line between them, but rather to be of a transition zone nature in reality.

The ductile–brittle criterion is discussed and illustrated at length in Chapter 8. This ductile–brittle behavior can also be viewed for a particular stress state as showing the change as a function of the T/C variation. For example, for uniaxial tension it is found that

$$\frac{T}{C} > \frac{1}{2} \quad \textit{Ductile}$$
$$\frac{T}{C} < \frac{1}{2} \quad \textit{Brittle}$$

while for simple shear the result is

$$\frac{T}{C} > \frac{1}{3} \quad \textit{Ductile}$$
$$\frac{T}{C} < \frac{1}{3} \quad \textit{Brittle}$$

All stress states have T/C designated regions of ductile or brittle behavior. Some particular stress states are entirely of the ductile type for all values of T/C while some are entirely brittle.

The limiting case of the yield/failure paraboloid at $T/C = 1$ is given by the Mises cylinder, Mises [5.1], Fig. 5.7. The ductile–brittle criterion divides the Mises cylinder into ductile-versus-brittle failure regions as shown. Most of the common stress states are on the ductile side of the division, but if the mean normal stress part of the stress state is sufficiently great, brittle failure will occur. This might seem to be a surprising result, but on further consideration it is to be expected. First consider the effect of temperature. Temperature variation can determine whether a given material fails in a ductile or a brittle manner. So too can pressure control the ductile-versus-brittle failure character. Sufficient pressure can convert a nominally brittle material into a

ductile material. Conversely, negative pressure can convert a nominally ductile material into a brittle material, as shown in Fig. 5.7.

The development and consequences of this comprehensive yield/failure theory for isotropic materials have been examined in many different ways, and will be taken up throughout this book. For applications where both yield strength and brittle failure are viewed together as generalized failure, the failure criteria (5.1) and (5.2) comprise the entire specification needed for analysis.

5.4 *T* and *C* versus *S* and *D*

If, as argued here, there exists a comprehensive theory of failure for isotropic materials, then the most favorable situation would be that of two calibrating failure properties in coordination with the two elastic properties of physical/mechanical behavior. The most common sets of elastic properties are either μ and k or E and ν. The present theory of isotropic material failure employs the uniaxial strengths T and C as associated with E and ν. But this suggests that one could as well take the shear strength, S, and the positive dilatational strength, D, as associated with μ and k. This alternative approach is a matter worth examining.

The fact that all stress states can be built up from the superpositions of states of distortion and dilatation could lead one to suspect that S and D may be more fundamental than T and C. Furthermore, this superposition capability suggests (but does not prove) that two failure properties should indeed suffice for describing all isotropic materials failures. Proof can follow only from experimental verification. Note that the symbols S and D used here are a little different from the like symbols used in Section 5.3, but they have the same general implications.

Next, the failure criterion used here and based upon T and C will be recalled. Then it will be expressed in terms of S and D to determine whether the formulation becomes more clarified, less clarified, or completely opaque compared with that using T and C.

The failure criterion (criteria) developed and used here are represented by (5.1) and (5.2). Take S as the failure stress in shear, and D as the failure stress in positive dilatation. Henceforth, D will be referred to as dilation failure. Then (5.1) becomes

$$\frac{\sigma_{\mathrm{ii}}}{3\mathrm{D}} + \frac{1}{6\mathrm{S}^2}\left[(\sigma_1-\sigma_2)^2+(\sigma_2-\sigma_3)^2+(\sigma_3-\sigma_1)^2\right] \le 1 \qquad (5.3)$$

where

$$\mathrm{S}^2 = \frac{\mathrm{TC}}{3}$$
$$\mathrm{D} = \frac{\mathrm{TC}}{3\,(\mathrm{C}-\mathrm{T})} \tag{5.4}$$

The fracture criterion (5.2) becomes

$$\sigma_1 \leq \frac{\mathrm{S}^2}{2\mathrm{D}}\left(\sqrt{1+12\frac{\mathrm{D}^2}{\mathrm{S}^2}}-1\right) \quad \text{if} \quad \frac{\mathrm{D}}{\mathrm{S}} \leq \sqrt{\frac{2}{3}} \tag{5.5}$$

which is expressed in terms of the largest principal stress.

The first observation to be made is that the fracture criterion (5.2), when expressed in terms of S and D, (5.5), is completely devoid of physical interpretation. If the comparison of using T and C versus S and D involved only the polynomial-invariants criterion there would be little to choose between them. But the real problem is with the fracture criterion. The fracture criterion is very important, and not just in applications. It admits the limiting case of perfectly brittle behavior, without which there is no closure to the theory, and no actual assurance of generality.

Consider the relationships of T and C with E and ν, and correspondingly that of S and D with μ and k. From an operational, linear elasticity point of view, knowing the properties μ and k is absolutely equivalent to knowing E and ν. But from a physical point of view and especially with implications for failure, the two sets of elastic properties may or may not be entirely equivalent.

The possibly deeper meaning of E and ν versus μ and k is that the scale of the load-bearing resistance or capability in the one case is concentrated into one property, E, rather than into two properties, μ and k. The Poisson's ratio ν serves a completely different role, as will be explored in Chapter 14.

The limiting cases for the two elastic property forms are specified by

E and ν

(i) $\nu = 1/2$
(ii) $\nu = -1$

no restriction on E (positive values)

μ and k

(i) $k/\mu \to \infty$ or $\mu/k \to 0$
(ii) $k/\mu \to 0$ or $\mu/k \to \infty$

Some formulations with μ and k involve indeterminacies and discontinuities, whereas those with E and ν usually do not. A non-trivial example will be given that illustrates this.

An evaluation example is needed that not only involves stress and strain under elasticity conditions, but also the example should have implications or significance for failure behavior. Since isotropy is being considered here it must comply with that as well. The ideal example turns out to be that for the effective behavior of an elastic medium having a randomly oriented and distributed array of cracks.

In a major work, Budiansky and O'Connell [5.2] presented the solution for the effective properties of an isotropic, elastic material containing a damage pattern of randomly oriented cracks. This problem and solution is ideal for present purposes. It brings in elastic properties, damage, and failure significance. The only potential problem with using this work is that some have questioned the meaning of its prediction of material collapse at a value of $\varepsilon = 9/16$, where ε is the crack-density variable of circular cracks.

This result for material collapse is sometimes criticized as being arbitrary, or arbitrarily small. So, before using the results of this work, its validity will be reinforced by showing the logic and consistency of its prediction of the loss of material integrity at $\varepsilon = 9/16$.

For circular cracks the crack density is defined as

$$\varepsilon = \frac{n}{V}\left\langle a^3 \right\rangle \tag{5.6}$$

where a is the radius of the randomly distributed cracks, with $\langle a^3 \rangle$ being an average, and n the number of cracks existing within the volume V. For extension to more general cracks of convex shape, the crack density is taken as

$$\varepsilon = \frac{2n}{\pi V}\left\langle \frac{A^2}{\rho} \right\rangle \tag{5.7}$$

where A is the crack area and ρ its perimeter.

To verify the prediction of material collapse at $\varepsilon = 9/16$, consider different possible crack patterns at material collapse. The faces of some of the platonic solids will be taken to represent the fully developed crack patterns at collapse. The only platonic solids or combination of platonic solids that pack perfectly in 3-space are the cube and a 2:1 combination of tetrahedra and octahedra.

Dodecahedra do not pack in 3-space, although they come fairly close to doing so. The dihedral angle for a dodecahedron is $\theta = 116.°57$. If this angle were 120° there would be perfect packing. As an estimate for that obtained from the dodecahedron, take the number of sides of the single cell as twelve, recognizing that if there were perfect packing the appropriate number would be six, since each face would be shared with the neighboring cell. This then recognizes the imperfect packing, leaving a small amount of material as being left over to form a "skeleton" of remaining connected material at or near collapse. In solving for ε for the case of the dodecahedron the volume will be taken as that of the dodecahedron itself. Thus this estimate for ε can be viewed as an upper bound.

The respective values of ε for these three cellular cases of fully developed crack patterns are determined from (5.7) to be as shown in Fig. 5.8.

In terms of the crack pattern symmetries, the first two cases have cubic symmetry, while the third is isotropic. Based on these results it is completely reasonable and rational to expect that $\varepsilon = 9/16$ gives the minimum crack density value for material collapse due to an isotropic pattern of cracks, and correspondingly for a random distribution of cracks.

As a consequence of the above comparison, the results of Budiansky and O'Connell will be taken as the rigorous solution of the problem. These results will now be used for present purposes. As an aside, the

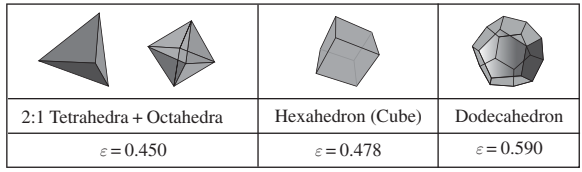


Fig. 5.8 *Self Consistent $\varepsilon = 9/16 = 0.5625$.*

generalized self-consistent method is the preferred effective properties approach for two-phase media, while the self-consistent method is the preferred approach for single-phase media containing cracks or grain boundaries. It is the latter method that was used by Budiansky and O'Connell.

The notation of Budiansky and O'Connell takes E_c, ν_c, μ_c, and k_c as the effective elastic properties of the medium containing the cracks. The same symbols without the subscripts are the properties of the uncracked material.

Of the four properties E_c, ν_c, μ_c, and k_c it is only ν_c as a function of ε that can be determined as an uncoupled equation apart from any other of the properties. Once ν_c is found, then the other three properties can be determined as functions of ε. In general, ν_c as a function of ε and ν is a cubic equation, but its analytical solution can be found in the special case of $\nu = 1/2$. This solution for circular cracks is given by

$$\nu_c = \frac{9}{16\varepsilon}\left[1 - \sqrt{1 - \frac{16\varepsilon}{9} + \left(\frac{16\varepsilon}{9}\right)^2}\right] \quad for \quad \nu = \frac{1}{2} \tag{5.8}$$

Eq. (5.8) is indeterminate at $\varepsilon = 0$, but it can be easily evaluated and gives the proper result of $\nu_c = 1/2$ at $\varepsilon = 0$. Knowing ν_c, the other properties can be found, such as for E_c, from

$$\frac{\mathrm{E}_c}{E} = 1 - \frac{16}{45}\frac{(1 - {\nu_c}^2)\,(10 - 3\nu_c)}{(2 - \nu_c)}\,\varepsilon \tag{5.9}$$

Similar forms are found for μ_c and k_c.

The general solutions for E_c, ν_c, μ_c, and k_c have the following characteristics. At $\nu = 0$,

$$\frac{\mathrm{E}_c}{E}, \frac{\nu_c}{\nu}, \frac{\mu_c}{\mu}, \ and \ \frac{k_c}{k}$$

all are the same linear function of ε going from the value of 1 at $\varepsilon = 0$ to the value of 0 at $\varepsilon = 9/16$. They gradually deviate from linearity as ν is increased from zero, going to maximum deviations at $\nu = 1/2$. Thus it suffices to examine the cases at $\nu = 1/2$ to see the maximum deviations from linearity.

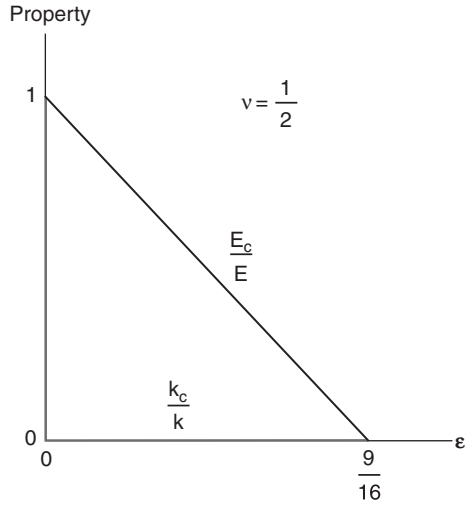


Fig. 5.9 *Properties variations as functions of crack density.*

The properties behaviors at $\nu = 1/2$ are as shown in Fig. 5.9. The slight deviation of E_c/E from linearity will be specified below, and k_c/k is discontinuous with ε. The variation of ν_c/ν in a plot of the type of Fig. 5.9 is a little above that for E_c/E, necessarily being concave down with the same end points, and that for μ_c/μ is considerably above that for ν_c/ν, also concave down. The deviations from linearity in plots of the type in Fig. 5.9 can be found as follows.

The mid-points of the variations, in Fig. 5.9, are at $\varepsilon = 9/32$. At this mid-value of ε and for this worst case of $\nu = 1/2$, the properties values are found to be

$$\begin{aligned} \frac{\nu_c}{\nu} &= 2\left(2 - \sqrt{3}\right) \\ \frac{\mathrm{E}_c}{E} &= \frac{6}{5}\left(1 - \frac{1}{\sqrt{3}}\right) \\ \frac{\mu_c}{\mu} &= \frac{3}{5} \end{aligned} \tag{5.10}$$

Thus the maximum deviations from linearity are found to be as given in Table 5.2.

Table 5.2 Property deviations from linearity

$\frac{E_c}{E}$	1.4% deviation from linearity
$\frac{\nu_c}{\nu}$	7.2% deviation from linearity
$\frac{\mu_c}{\mu}$	20% deviation from linearity
$\frac{k_c}{k}$	Discontinuous

The result that k_c be discontinuous when $\nu = 1/2$ can be easily reasoned as being physically required. Recognizing the intimate relationship of the effective properties to fracture mechanics, it is reasonable to expect that the dilation failure property, D, would also suffer a discontinuous or nearly discontinuous behavior with variations of ε.

It is seen that E_c and ν_c have a special behavior not shared by μ_c and k_c. In other sections it has been discussed that failure criteria can be interpreted as the tensorial description of the termination of the elastic range of behavior. In this sense there is an association of elastic properties and failure properties. Plastic flow behavior does not contradict this; it just describes a more complex transition from elastic behavior to failure. In relation to this, the behaviors of E_c and ν_c, as found by Budiansky and O'Connell [5.2] over the full range of crack densities, show vastly more regularity and smoothness than do μ_c and k_c. Furthermore, the difference in behavior in this problem between E_c and ν_c is minor, while the disparity between μ_c and k_c behaviors is huge.

The special behavior of the Poisson's ratio ν is quite interesting. Although T and C relate most directly to E itself, ν also may have some meaning for failure behavior. Most but not all engineering materials have Poisson's ratios that fall within $1/5 \leq \nu \leq 1/2$. Nothing of this type can be said for μ and k except by inference from ν. Furthermore, Poisson's ratios give some indication of a materials failure type. Within the range stated above, the lower Poisson's ratios relate to brittle failure, while the mid-range and some higher ones usually relate to ductile failure. A much deeper look at the significance of Poisson's ratio will transpire in Section 14.4.

In conclusion of these results and observations, the elastic properties E and ν and their failure conjugates T and C are much more likely to be the fundamental continuum mechanics pairings than are the elastic properties μ and k and their failure conjugates S and D. It may also be noted that as a practical expedient, the uniaxial strengths, T and C, are enormously more easy to experimentally determine with reliability than are the shear and dilation strengths, S and D. Agreeably then, the fundamental forms and the practical approach for failure properties determination are in harmony, not in conflict.

Problem Areas for Study

1. Do failure criteria have a place and a use for materials such as elastomers that have no effectively linear elastic range of behavior?
2. The existence of significant linear elastic ranges of behavior in engineering materials is an extremely fortunate but rather poorly understood phenomenon. For example, in ductile metals, dislocation emission occurs upon the first small increment of load application, but at some point the rate of emissions multiplies hugely and cascades into macroscopic yield/failure. Many other materials types show similarly sudden changes in behavior with load application. What are the underlying physical mechanisms that cause such abrupt changes in constitutive behavior? Can some or most cases be explained by localization, or is every different materials type explained by a different mechanism?
3. Typical failure criteria approaches focus upon a specific class of materials, such as ductile metals, brittle metals, polymers, and so on. They then fit parameters to data for specific stress states. Can such empirical approaches be taken to apply for other stress states if they do not apply across the various materials classes and thereby certify generality in that way?
4. There are many molecular level theories for amorphous or cross-linked, glassy polymers. Generally they involve many parameters to reflect a variety of features in molecular scale architecture and bonding. How can these be related usefully to a two-property macroscopic level failure criterion?
5. The two-property failure theory developed here for isotropy has its calibration from T and C, the two uniaxial strengths. Can a viable two-property theory be developed in terms of S and D, the shear strength

and the positive dilatational strength? It should be noted that the present T and C theory cannot be converted to an S and D form, because S and D forms cannot produce the fracture criterion present in the T and C theory. See Section 5.4. Does this preclude an S and D theory?

References

[5.1] Mohr, O. (1900). "Welche Unstande Bedingen die Elastizitasgrenze und den Bruch eines Materials," *Zeitschrift des Vereines Deutscher Ingenieure*, **44**, 1524–30.

[5.2] Budiansky, B. and O'Connell, R. J. (1976). "Elastic Moduli of a Cracked Solid," *Int. J. Solids Structures*, **12**, 81–97.

6
Experimental and Theoretical Evaluation

The term "evaluation" implies more than just a routine comparison of any particular isotropic material failure theory with some convenient data. First, a candidate failure criterion should undergo a rigorous examination of its physical and mathematical basis. Then it should be experimentally evaluated using only the highest-quality testing data from carefully selected, homogeneously applied stress states. Such formal and focused probes are necessary if realistic (and reliable) failure forms are to be successfully identified.

All of this is best preceded by a reading of the very long and highly unusual search for failure criteria for isotropic materials, such as is given in Chapter 2. The account here will begin with a brief, telescoped history and its consequent status in order to provide the proper background for the explicit evaluation.

6.1 Evaluation Problem

Perhaps the simplest imaginable failure criterion for all homogeneous and isotropic materials would be that of a critical value of the total stored energy in the material. Even though this notion was dispelled at least 150 years ago, it is frequently rediscovered and enthusiastically proposed.

One of several roadblocks for this simple energy idea is that the yielding of very ductile metals depends upon only the critical value of the distortional part of the total energy, not the total energy itself. This, of course, ties in with the fact that dilatational stress states cannot move dislocations, as only shear stresses can do so. However, a distortional energy criterion (Mises criterion) does not work for anything but a very ductile material, as is easily demonstrated. Therefore, something far more sophisticated than an energy criterion must be required.

After languishing for more than 100 years, the original Coulomb criterion was put into a very utilitarian form by Mohr. At the time, the resulting Coulomb–Mohr form was widely thought to have all the answers. That transpired a little less than 100 years ago, but it too turned out to be a false hope. Maximum normal stress and maximum shear stress criteria also proved to be disappointing.

In the years following the notable Coulomb–Mohr effort and over some very long passage of time, the search for failure criteria seemed to degenerate into essentially just postulating forms. These may have had analytical appeal to the originators, but they had little or no physical foundation for support and further development.

In simplest terms, stiffness and strength (and toughness) have independent properties. The underlying basic problem is to express these independencies in full tensor forms and to specify the explicit mechanical properties in each that control three-dimensional behavior. Of the following three areas of analytical mechanics characterization—stiffness, strength, and toughness—only that of strength still has unresolved problems after all this elapsed time. There is a broad lack of understanding with respect to the three-dimensional characterization of failure conditions—that is, failure criteria. Consequently this technical area remains in a state of uncertainty and disorganization. The other two areas—stress–strain relations (stiffness) and fracture (toughness)—are extremely highly developed.

With the historical results for failure criteria appearing to be so meager and unproductive, there is a legitimate question as to whether the problem may be too difficult to ever yield a general solution. Some would say it will always be approachable only on an empirical case-by-case basis.

There is, however, a completely contrary and opposing point of view. This view sees the historical efforts as having provided a promising "foothold" on the problem, to be used to advantage when the opportunity should arise. In particular, the Mises criterion, despite its restrictions, could serve very well to anchor a more general approach.

Rather than treating failure criteria as merely an adjunct to the stress–strain relations, the two should be recognized as being of completely different character, thereby requiring absolutely independent constitutive developments. That is the direction followed here and in supporting works. The technical derivation of Chapter 4 will be summarized, followed by the evaluations of the resulting failure criteria using specific sets of data.

The emphasis here is on isotropic materials. This is not just because they are so overwhelmingly important, but also because they produce the guidelines for all other cases. However, anisotropic, composite materials will also be given due attention later.

6.2 Theoretical Assessment

Failure criteria of the type considered here are not needed for all materials. For example, elastomers and "soft" biological tissues probably have no need for explicit three-dimensional failure criteria. On the other hand, for engineering materials with essential stiffness characteristics, there are limits to this performance and failure criteria are of vital importance in applications. Mainly those materials that possess a range of linear elastic behavior (or nearly so) are those that require the characterization of the limits of that performance. These include metals, the glassy (but not rubbery) polymers, some wood products, ceramics, glasses, materials of construction, many geological materials, fiber composite materials, and many others as well. Most of these materials are isotropic or nearly isotropic.

The requisite characteristic is that of an effectively linear elastic behavior up to, but not beyond, some load level. Failure and linear elasticity share a special type of inverse and complementary relationship. The ideal linear elastic behavior is terminated by the failure, and all the non-linearity is subsumed into the failure criterion. Historically, linear elasticity theory was the progenitor of all classical field theories, but the related failure theory is required to complete the characterization.

A failure criterion takes the form of a surface in stress space. The stress space can either be the 3-space of the principal stresses, or the larger space of the full stress tensor, or smaller subspaces. The comparison with geometry is obvious. A geometric surface in 3-space is described by a scalar function of the coordinates $f(x_i) = 1$. In the stress context, the problem is to find the scalar equivalent for the failure criterion problem, $F(\sigma_{ij}) = 1$, giving the failure surface in stress space.

So the problem distills to finding the proper form for the scalar function—here called the scalar potential of failure. The failure potential is not to be confused with the plastic potential, which has a different meaning and use.

The failure criterion derivation of Section 4.4 will be briefly summarized here in a slightly different form, so that it can be compared with another failure criterion in the overall theoretical evaluation process. The form of the failure potential will be taken from the same mathematical representation as that used with the elastic energy potential, since it too is a scalar and in some broad sense related. However, significant and formative differences between the two will arise immediately thereafter. As with elastic energy, take the failure potential as a polynomial expansion in the invariants of the stress tensor, as appropriate to isotropy. This then is the method of polynomial invariants, giving

$$\phi = a_0 + a_1 I_1 + a_2 I_1^2 + a_3 J_2 + - - - \tag{6.1}$$

I_1 is the first invariant of the stress tensor, and J_2 the second invariant of the deviatoric stress tensor, as

$$\begin{aligned} I_1 &= \sigma_1 + \sigma_2 + \sigma_3 \\ J_2 &= \frac{1}{2}\left[(\sigma_1 - \sigma_2)^2 + (\sigma_2 - \sigma_3)^2 + (\sigma_3 - \sigma_1)^2\right] \end{aligned} \tag{6.2}$$

where principal stresses are used. See Chapter 4 for the more general case not involving principal stresses. The expansion is terminated at terms of second degree—the same as with the elastic energy representation. Parameter a_0 merely establishes the datum and is not needed here. Sometimes a coefficient different from 1/2 is used in the J_2 expression, which is acceptable only so long as it is used consistently.

In the case of the elastic energy potential the resultant form must be positive definite and therefore $a_1 = 0$, but for the failure potential there is no such requirement and so $a_1 \neq 0$. However, for a physical condition that the homogeneous and isotropic material not fail under hydrostatic compression, it is required that $a_2 = 0$, leaving the failure potential as

$$\phi = a_1 I_1 + a_3 J_2 \tag{6.3}$$

Then the failure criterion becomes

$$a_1 I_1 + a_3 J_2 \leq 1 \tag{6.4}$$

Finally, evaluating a_1 and a_3 in (6.4) in terms of the uniaxial tensile and compressive strengths, T and C, gives

$$\left(\frac{1}{T}-\frac{1}{C}\right)I_1+\frac{1}{TC}J_2\leq 1 \tag{6.5a}$$

or

$$\begin{aligned}&\left(\frac{1}{T}-\frac{1}{C}\right)(\sigma_1+\sigma_2+\sigma_3)\\&\quad+\frac{1}{2TC}\left[(\sigma_1-\sigma_2)^2+(\sigma_2-\sigma_3)^2+(\sigma_3-\sigma_1)^2\right]\leq 1\end{aligned} \tag{6.5b}$$

As the overall failure criterion, (6.5) is necessary but not sufficient. Sufficiency requires that a competitive fracture criterion in the more brittle range of behavior also be satisfied. From Section 4.5 this is given by

$$\left.\begin{aligned}\sigma_1&\leq T\\ \sigma_2&\leq T\\ \sigma_3&\leq T\end{aligned}\right\}\quad \textit{if}\quad \frac{T}{C}<\frac{1}{2} \tag{6.6}$$

There is both geometric and physical significance to the commencement of the fracture criterion at $T/C = 1/2$. It is at this value of T/C that the three fracture planes (6.6) in principal stress space are just tangent to the paraboloid (6.5) of the failure potential. The "if" condition in (6.6) can be written either as the inequality shown or as an inequality with equality because of the tangency condition. The fracture criteria (6.6) then introduces the occurrence of "corners" into the yield/failure surface characterization for $T/C < 1/2$.

The forms (6.5) and (6.6) comprise the complete, two-property failure criterion for homogeneous and isotropic materials. In effect, the failure criterion and the elastic energy (from which the stress–strain relations follow) are joint constitutive relations. They are independent, but one without the other is incomplete. The elastic energy has two properties (E and ν) and one condition (positive definiteness), while the failure potential also has two properties (T and C) and one condition (no failure under hydrostatic compression).

Placing the above results aside until experimental evaluation in the next section, a contemporary example of a postulated failure criterion will now be given as a case study in contrasts. A research group has

for several years advocated that the material failure criterion (for the isotropic matrix phase in composites) be written directly in uncoupled, two-parameter form as

$$\begin{aligned} I_1 &\leq \alpha \\ J_2 &\leq \beta \end{aligned} \tag{6.7}$$

The first of these is said to associate with a fracture behavior, and the second with yielding and flow. Taking both α and β to be calibrated by both T and C then gives

$$\begin{aligned} I_1 &\leq T \\ J_2 &\leq C^2 \end{aligned} \tag{6.8}$$

It is not clear what the physical basis of (6.7) and (6.8) may be. The equations merely reflect that J_2 applies to ductile metals, and there are also (erroneous) claims that I_1 completely covers brittle materials, so the proponents take both of them as independently and simultaneously required for all glassy polymers (and by extension for all materials). This is typical of the many schemes that have been tried. Certainly all obvious ways of combining the invariants (including I_3 or J_3) have been proposed and in some cases claimed to be the long-sought form for failure criteria. Most of these attempts do not receive a critical comparison with data on well-characterized homogeneous and isotropic materials. Rather, they usually involve comparisons with a window of data from non-uniform stress states in complicated forms or structures where many other factors are also in effect, but not acknowledged nor even recognized.

With regard to the possibility of forms (6.8) being a useful, well-posed failure criterion, consider the following. One of the basic tenets of materials science is that superimposed pressure increases the strength of the material. This happens in all cases except the limiting case of a Mises material. The failure criteria derived here, (6.5) and (6.6), embody this effect in an interactive manner. In contrast, the forms (6.8) predict that pressure has no effect at all on the uniaxial compressive strength nor on the shear strength, while they provide an unusual prediction for the effect of pressure on the uniaxial tensile strength.

Even more objectionable is the prediction of (6.8) in the case of the ductile limit at $T = C$. For two-dimensional biaxial stresses, (6.8) gives only a portion of the usual Mises ellipse because the first of (6.8) truncates

it by a straight-line cutoff in the first quadrant. In the range of polymer behavior, (6.8) still exhibits the same characteristic. At $T = C$ the derived failure criteria (6.5) and (6.6) directly recover the full Mises criterion.

All of these difficulties with forms (6.7) and (6.8) still occur when the invariants are interpreted in terms of strains rather than stresses. This is easily shown by substituting the isotropic stress–strain relations into (6.7), whichever way they are specified, to directly obtain the other form. Thus (6.7), whether in their stress or strain forms, fail to pass standard tests of physical consistency. All tentative failure criteria should be subjected to conceptual tests of the types just discussed. The cloud of ill-conceived failure criteria over the years has done much to obscure and impede progress.

There also have been a few very clever constructs, such as the Coulomb–Mohr form, but unfortunately they do not coordinate at all well with the physics of failure. This will be further shown through comparisons with data in the next section, since the Coulomb–Mohr criterion is by far the most substantial and prominent of all historical forms.

6.3 Experimental Evaluation

We return now to the isotropic material failure criteria (6.5) and (6.6). As a conjecture, the form (6.5) has been known for a very long time—at least from about the time of Mises. The derivation of it presented in this book adds substance and context, revealing its dual relationship to energy. It is, however, the coordination of (6.5) and (6.6) that provides the comprehensive core extending from ductile limit to brittle limit. Nevertheless, none of this is of any consequence unless it reflects physical reality. The criteria (6.5) and (6.6) must now pass tests of evaluation with meaningful data, or be rejected.

If the objective were merely to characterize failure for a single material type—epoxy polymers for example—then that could not really be considered as developing failure criteria. In reality it would just be "curve fitting" for that material with all the hazards and pitfalls for anything but the simplest interpolations. To validate a failure criterion it is necessary to test it against widely different classes of materials. Otherwise there would be no confidence in its general applicability. The qualifier "general applicability" refers not only to different types of materials, but also to all possible stress states for each.

More specifically, for isotropic materials it is necessary to cut across the full range from very ductile to very brittle materials. The extremes are quite apparent. A very ductile metal should be used to test the one extreme, and most likely a very brittle geological material should be tested for the other extreme.

The intermediate case is somewhat more problematic. However, at least two possibilities are at hand for a material that is neither very ductile nor very brittle. These two would be cast iron and very glassy, untoughened polymers. Of the two, there appears to be more high-quality data for cast iron, and it is what will be included here. Although cast iron is loosely thought of as being brittle, that is only so in comparison with ductile steels. Over the full range of ductile/brittle behavior, cast iron lies as ideally intermediate for this purpose. This will be explained more fully later.

Before looking at individual data cases, one more testing requirement should be observed. It is very important that testing results not come from conditions containing steep stress gradients. Only homogeneous or nearly homogeneous stress states provide the clarity level consistent with seeking strength-type mechanical properties for homogeneous materials.

The initial test case concerns the ductile limit at $T = C$. There are several different metals that could be used to generate data. The first substantial and enduring test results were those of Taylor and Quinney [6.1] in 1931. They used the tension–torsion of thin-walled tubes to generate plots of normal stress σ_{11} versus shear stress σ_{12}. The two theoretical forms of interest are the Mises criterion and the Tresca criterion. The Mises criterion is the form from the general failure criteria (6.5) and (6.6) at $T = C$, and it also represents a distortional energy criterion. The Tresca criterion is that of the Coloumb–Mohr theory at $T = C$, and it also represents a maximum shear stress criterion.

The comparison between the two theories and the ductile metals data are as shown in Fig. 6.1. The data clearly favor the Mises criterion and thereby the general criteria (6.5) and (6.6). G. I. Taylor was one of the great scientists of the twentieth century. These results from him have been substantiated many times and are now considered to be classical. R. Hill [6.2], who was also a world-class contributor, affirmed this status for these results.

Next consider the intermediate case that is between extreme ductility and extreme brittleness. This case is that of a particular type of

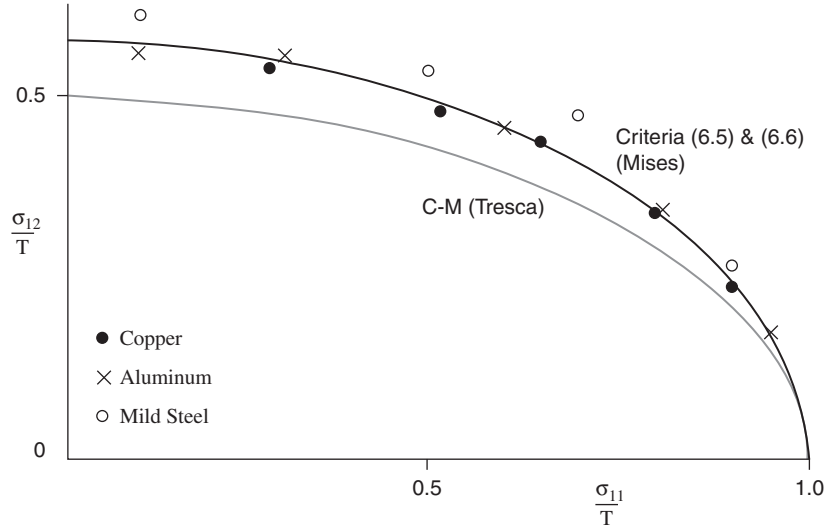


Fig. 6.1 *Taylor and Quinney [6.1] failure data for ductile metals, $T/C = 1$.*

cast iron. In such cases the T/C ratio is about 1/2 to 1/2.5, which is effectively in the middle range between $T/C = 1$ for ductile metals, and T/C of about 1/15 or even much less for brittle geological materials. The biaxial cast iron failure data in Fig. 6.2 are by Cornet and Grassi [6.3], the first of whom was a careful experimentalist at University of California at Berkeley. The T/C value is 1/2.16.

The comparison of the data in Fig. 6.2 with failure criteria (6.5) and (6.6) is favorable, while the comparison with Coulomb–Mohr is unfavorable. These data have been effectively reproduced by several other investigators for other types of cast iron. The fracture criterion (6.6) is just barely discernable as a slightly flattened "spot" on the failure envelope of Fig. 6.2. The Mises or Tresca criteria by themselves would be completely unsatisfactory in this middle-range case.

As the third critical test case, a very brittle material must be used. Reliable failure data were generated on dolomite by a respected geophysicist at the Massachusetts Institute of Technology. These results from Brace [6.4] are on samples considered to be close to isotropic, and with $T/C = 1/14.9$.

The failure criteria (6.5) and (6.6) provides a good correlation with the data in Fig. 6.3, whereas Coulomb–Mohr fails to follow the downturn in the data and merely projects a straight-line failure envelope.

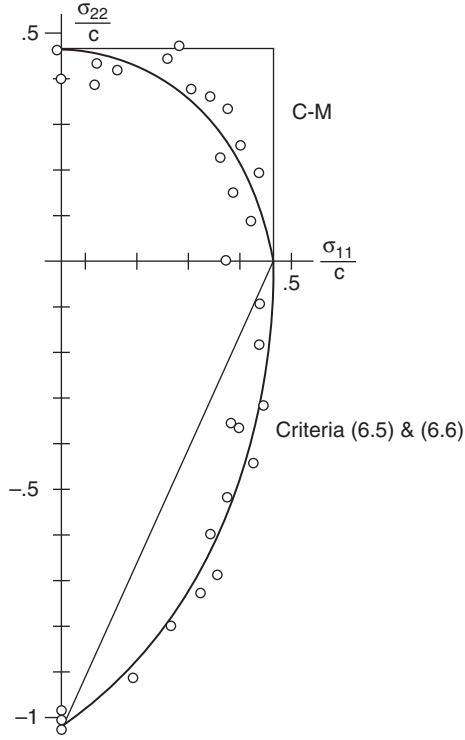


Fig. 6.2 *Cornet and Grassi [6.3] biaxial failure data on iron, T/C = 1/2.2.*

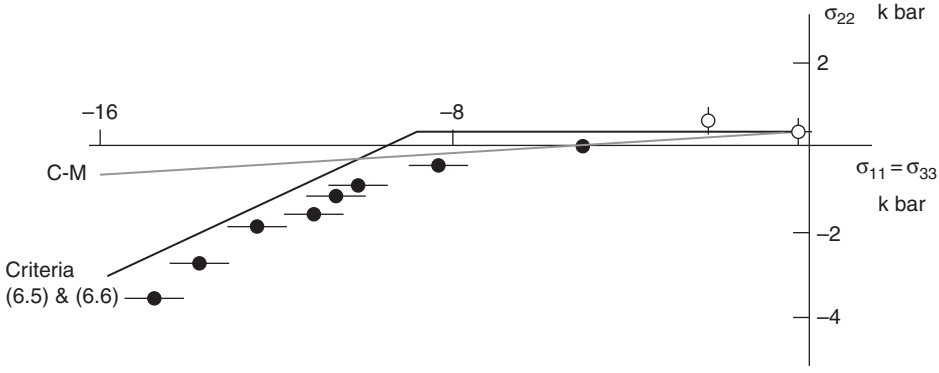


Fig. 6.3 *Brace [6.4] tri-axial failure data for Blair dolomite, T/C = 1/15.*

The upper envelope shown in Fig. 6.3 is that due to the fracture criterion (6.6). The experimental uncertainty along the horizontal axis is quite considerable and within the proximity of the predicted failure envelope. Probably the most difficult (and uncertain) testing result is that for uniaxial compressive strength C. If the value of C given in [6.4] were to be reduced by 10% the prediction from (6.5) for the downturn region would fall within the data scatter.

If the stress states were taken with σ_{11} and $\sigma_{22} = \sigma_{33}$ the theoretical envelope would have the downturn just inside the data, rather than just outside the data, as in Fig. 6.3. If one extends the axes in Fig. 6.3 one sees the tremendous divergence between the Coulomb–Mohr theory on the one hand and the data and the present theory on the other hand. The downturn in the data and theory shown in Fig. 6.3 is not surprising. The same effect shows up in the brittle limit discussed in Chapter 5. It should be noted that the axes labels in Fig. 6.3 are different from those used by Brace in order to follow common convention.

A comprehensive failure theory must span failure mode types from ductile flow to completely brittle disintegration. It is nearly inconceivable that this could be accomplished without a formal and rigorous framework for treating ductile-versus-brittle behaviors. The present failure theory, which culminates in the explicit criteria (6.5) and (6.6), does have an underlying ductile/brittle formalism. This was mentioned in Chapter 5 and is presented in Chapter 8. The complete formalism gives not only the failure envelopes, but also the division of domains into ductile-versus-brittle regions. This will be briefly summarized here, since it too is amenable to experimental verification or repudiation.

The concept of the ductile/brittle transition is basic to materials science. Usually this is posed as a temperature effect, and the ductile/brittle transition temperature can be located by scanning failure type against temperature change. The same approach can be taken with a preimposed state of pressure. Sufficient pressure can convert a perceived brittle material into an effectively ductile material, and there is a transition from one state to the other.

This interpretation of the ductile/brittle transition admits extension to a variety of other conditions. For example, any particular state of three-dimensional stress dictates whether the particular failure mode is ductile or brittle. For a given material, predominantly tensile stress states are far more likely to be brittle in failure than are predominantly compressive

stress states. Whether a particular failure mode will be ductile or brittle depends explicitly and crucially upon the stress state itself, as well as the material. Conversely, for a particular stress state one can scan across various materials types to determine the ductile-versus-brittle nature of the failure mode.

Following the above prescription, for any particular stress state the material type will be varied by varying the T/C ratio, and then the failure mode type, ductile or brittle, will be identified. From the rather lengthy derivation in Chapter 8, the ductile/brittle transition follows from the intersection of failure criteria (6.5) and (6.6) as

$$\begin{aligned} I_1 &= 3T - C \qquad\qquad D/B \;\; \textit{Transition} \\ &= (T + C) + 2(T - C) \end{aligned} \tag{6.9}$$

where $I_1 = \sigma_{11} + \sigma_{22} + \sigma_{33}$ is the first invariant for the stress state of interest that is on the failure envelope from (6.5). The ductile and brittle conditions are then given by

$$\begin{aligned} I_1 &< 3T - C \qquad \textit{Ductile} \\ I_1 &> 3T - C \qquad \textit{Brittle} \end{aligned} \tag{6.10}$$

where again I_1 is from (6.5). Relation (6.9) is physically meaningful only if the solution for T/C falls in the range from 0 to 1. Otherwise (6.10) shows the condition to always be ductile or brittle for all allowable values of T/C.

It would be expected that the criterion for the ductile/brittle transition would involve the first invariant part of the total stress state at failure. This can be thought of as algebraic pressure and that has an intimate relationship with temperature in changing the state of substances.

A simple example will illustrate the use of (6.9) and (6.10). Take the state of uniaxial tension. For this case the solution from (6.5) is almost trivial—it is $I_1 = T$—and substituting that into (6.9) gives

$$\frac{T}{C} = \frac{1}{2} \quad \textit{Uniaxial Tensile } D/B \textit{ Transition}$$

For materials with $T/C > 1/2$ the tensile failure is predicted by (6.10) to be ductile, and for $T/C < 1/2$ the tensile failure is predicted as

brittle. This is in accordance with most experimental observations for a wide variety of different materials types. Most metals and most polymers are found to be on the ductile side of the predicted D/B transition at $T/C = 1/2$, while a few metals, a few polymers, and all ceramics, glasses, and geological materials are on the brittle side. The three preceding data test cases are in agreement with this classification of behavior. The complete spectrum of ductile/brittle uniaxial tensile behaviors is yet another aspect of experimental corroboration of the general theory.

For uniaxial compression, from (6.5) $I_1 = -C$ and that into (6.9) gives $T/C = 0$. Thus uniaxial compression is always ductile, but becomes borderline brittle as T/C approaches the brittle limit.

The range covered by this failure theory is $0 \leq T/C \leq 1$. An interesting sequence of quite simple stress states that covers the full range of ductile/brittle T/Cs is given in Table 6.1.

It follows from (6.10) that equal triaxial tension is brittle for all values of T/C, and equal biaxial compression is ductile for all values of T/C. It also is recalled that there is no failure under equal triaxial compression.

The two uniaxial stress states have already been discussed. The state of simple shear is also of importance. For simple shear $I_1 = 0$, and then (6.9) gives the ductile/brittle transition as being at $T/C = 1/3$. From the failure criteria (6.5) and (6.6) at $T/C = 1/3$ it is found that the shear failure stress is exactly the same as that of uniaxial tension at this value of T/C. Shear stress rotated 45 degrees is equivalent to a tensile stress component and an orthogonal compressive stress component of the same magnitude. This shows that the compressive stress component has

Table 6.1 Ductile/brittle behavior for different stress states

Stress State	D/B transition $T/C =$	D/B behavior T/C
Equal biaxial tension	1	<1 brittle
Uniaxial tension	1/2	>1/2 ductile <1/2 brittle
Simple shear	1/3	>1/3 ductile <1/3 brittle
Uniaxial compression	0	>0 ductile

no effect on the failure mode; it is entirely caused by the tensile stress component. This is symptomatic of brittle behavior.

Perhaps the most interesting case in Table 6.1 is that of the equal biaxial tension. The solution from (6.5) for $\sigma_1 = \sigma_2$ and $\sigma_3 = 0$ gives the tensile root as

$$I_1 = 2\left(T - C + \sqrt{T^2 - TC + C^2}\right)$$

and this into (6.9) gives the D/B transition as being at $T/C = 1$. Now, a material such as aluminum or steel with $T/C = 1$ is certainly thought of as being a ductile material, and it most assuredly is ductile in uniaxial tension. But the result in Table 6.1 shows that such normally and nominally ductile materials but in a state of equal biaxial tension are actually at the transition between being ductile and brittle.

Experimental pressurization of ductile steel spherical pressure vessels reveals failure modes that are more suggestive of brittle behavior than of ductile behavior. Typical failure modes for this situation have been shown and discussed by Talja *et al.* [6.5]. The failures of spherical steel pressure vessels are explosively sudden. There is no evidence of large-scale plastic deformation. Even though the material is thought of as being highly ductile, the failure mode type is that of fracture, causing either the formation of large fragments, or blowing out plugs, as in [6.5]. All this is consistent with the failure type being borderline brittle for the state of equal biaxial tension, as discussed above.

In principal stress space the failure criterion (6.5) takes the form of a paraboloid, while the fracture criteria (6.6) are planes that cut off portions of the paraboloid. The ductile/brittle transition specification (6.9) is that of a plane normal to the axis of symmetry of the paraboloid. It divides the paraboloid into ductile and brittle regions. For $T/C < 1/2$ it also cuts across the three fracture planes, thus giving failure modes of brittle fracture and of ductile fracture types.

6.4 Isotropy Conclusion

There are many other isotropic materials failure criteria that have been proposed and (inappropriately) used. These can easily be evaluated using the data cases just discussed. For the Mises and Tresca criteria, only the Mises form is completely satisfactory for ductile metals, but both are

in serious error for everything else. Of the two-parameter variety, these include the maximum stress (tensile and compressive), the maximum strain, and the Drucker–Prager forms. All these compare very poorly with the data cases. Equally importantly, all the above criteria fare no better in assessing their theoretical roots. The inadequacy of the Coulomb–Mohr form has already been shown here many times.

Other data evaluations for other materials types including polymers and ceramics are given in [4.1]. These results are compatible with the conclusions reached in this evaluation chapter.

Three-parameter forms have not been considered here because of the success of the two-property theory that has been the main focus. In recent times, three-parameter forms have been used mainly in problems of crack formation in ductile metals after experiencing extreme plastic deformation. Typically, the third parameter involves the third invariant I_3, often in the form of the Lode angle.

The failure theory developed here is for a totally different purpose than that just mentioned. The present work is relevant to the onset of dominating irreversible deformation, whether it be due to the inception of major plastic flow, as a form of failure or explicit, brittle failure itself, or due to any other mechanism.

The overall conclusion from this failure theory and its experimental evaluation is that the two properties T and C in combination provide a fundamental and comprehensive calibration of materials strength performance. The defining two-part failure criterion (6.5) and (6.6) embedding these two properties completes the constitutive specification for homogeneous and isotropic materials. Of course, this ambient-conditions theory would not be expected to apply under extreme pressure (and temperature) conditions where the magnitude of the first invariant at failure is much larger than the compressive strength C.

Problem Areas for Study

1. The effect of superimposed pressure upon failure is in need of a comprehensive and critical treatment for isotropic materials. The present failure theory predicts that shear stress with pressure gives the failure envelope as shown in Fig. 7.3 in the next chapter. Other stress states admit similar forms. Predictions such as these should be compared with the famous Bridgeman data. In so doing, the efficacy of this quite old

data must be assessed. Follow this line of investigation and supplement it with newer sources of data wherever possible.

2. There is a great need for critical experiments that test failure criteria in new and original ways. Difficult and expensive though such testing may be, it must remain as a prime objective. What should be the most important guidelines and objectives to be investigated in the testing?
3. In comprehensively evaluating failure criteria by comparing with data, would success with cases near the ductile limit and also near the brittle limit assure success across the full spectrum of materials types?
4. How many different types of stress states should be sought for testing verification? Not all stress states have equal sensitivity. For example, for failure criteria calibrated by T and C, when tested using the shear strength S, there is little difference in predictions from the different theories.
5. Can failure surface orientations be used to evaluate failure criteria? Some work on this approach has already been done. Obviously, when a particular failure theory makes an egregiously incorrect prediction of a failure surface orientation for some particular materials type in some particular stress state, then it should be disqualified from further consideration. More broadly, however, would this method of evaluation be expected to give sensitive, general, and definitive results?
6. It is important to use both theoretical means and experimental data to evaluate failure criteria. Considering how difficult it is to produce reliable and sensitive experimental data, how should one balance the data significance versus the theoretical foundation significance?
7. In evaluating failure criteria from different test states, it is preferable that the stress state be of homogeneous character in the test specimen. Does this mean that tests with inhomogeneous stress states cannot be used for evaluation purposes? How should one draw the line between the two types?

References

[6.1] Taylor, G. I. and Quinney, H. (1931). "The Plastic Distortion of Metals," *Phil. Trans. Roy. Soc. London*, **A230**, 323–62.

[6.2] Hill, R. (1998). *The Mathematical Theory of Plasticity*, Oxford University Press, Oxford.

[6.3] Cornet, I. and Grassi, R. C. (1955). "Fracture of Innoculated Iron under Biaxial Stress," *J. Appl. Mech.*, **22**, 172–4.

[6.4] Brace, W. F. (1964). "Brittle Fracture of Rocks," in *State of Stress in the Earth's Crust* (edited by W. R. Judd), American Elsevier, New York.

[6.5] Talja, H., Keinanen, H., Hosia, E., Pankakoski, P. H., and Rahka, K. (2003). "Rupture Tests with Reactor Pressure Vessel Head Models," *Trans. 17th Int. Conf. Structural Mech. in Reactor Tech.* (SMiRT–17), Prague.

7
Failure Theory Applications

Some examples of applications will now be presented. The purpose is not to evaluate and compare with data, as that was covered in Chapter 6. Rather, the intention is to show the extraordinarily broad range and vitality of applications for the failure theory of Chapter 4.

All ductile metals are at or very near the strength ratio $T/C = 1$, but all other materials classes such as polymers, ceramics, glasses, and geological materials have broad ranges of T/Cs. Examples from the various materials classes will be used with particular but realistic values of T/C in each case. These examples will range from very ductile polymers to extremely brittle geological materials. There is no need to cover ductile metals in the examples, as they have been exhaustively treated in many places for a very long time.

The Coulomb–Mohr failure theory was completely discredited by von Karman and by Böker not long after Mohr finished its development. That situation has been revisited and re-evaluated in Section 3.5, with the same end result. Still, it is useful to show Coulomb–Mohr results here in the examples in order to compare it with and calibrate the new, comprehensive failure theory.

The general, isotropic material failure criteria will be stated in principal stress space. The principal stresses will be taken to be ordered as

$$\sigma_1 \geq \sigma_2 \geq \sigma_3$$

The failure theory treated here is given by two coordinated but independent criteria:

Polynomial-invariants failure criterion

$$\left(\frac{1}{T} - \frac{1}{C}\right)(\sigma_1 + \sigma_2 + \sigma_3) + \frac{1}{2TC}\left[(\sigma_1 - \sigma_2)^2 + (\sigma_2 - \sigma_3)^2 + (\sigma_3 - \sigma_1)^2\right] \leq 1$$

and

Coordinated fracture criterion

$$If \;\; \frac{T}{C} \leq \frac{1}{2} \;\; then \;\; \sigma_1 \leq T$$

The coordination refers to the physical basis for the inception of the fracture criterion at $T/C = 1/2$. This is a necessary and integral part of the failure theory; see Chapter 4. The two failure mechanisms are competitive, and whichever one produces the lesser allowable stress is controlling. In the examples these two criteria will be referred to as the failure criterion and fracture criterion respectively.

The Coulomb–Mohr (Mohr–Coulomb) criterion is:

Coulomb–Mohr failure criterion

$$\frac{\sigma_1}{T} - \frac{\sigma_3}{C} \leq 1$$

In the examples this will be referred to as the C–M criterion.

7.1 Very Ductile Polymers

The problem is that of the strength behavior of a very ductile, semi-crystalline polymer, polypropylene, under superimposed pressure. Polypropylene (and polyethylene) are very-large-volume commodity plastics with a large variety of uses and applications. Although these materials have rather low stiffness and strength, it is their ductility and toughness that enables most applications. In the present example the uniaxial strengths are from Pai [7.1] at

$T = 37.2$ MPa
$C = 45.5$ MPa

The strength ratio $T/C = 0.82$ denotes a very ductile behavior. This then could suggest that the Mises criterion would suffice for describing the failure characteristics. The main purpose of this example is to show just how wrong that casual assumption would be. For the material under pressure p the governing failure criteria are given by the following.

Table 7.1 Ductile polymer strengths under pressure

	σ_{11}^{T}	σ_{11}^{C}	S
Data	61.4 MPa	–74.5	40.0
Failure criterion	60.9	–69.2	37.5
Mises criterion	37.2	–37.2	21.5

Failure criterion

$$\sigma_{11} = \frac{1}{2}\left[T - C \pm \sqrt{(T+C)^2 + 12\,(C-T)\,p}\,\right]$$

$$S = \sqrt{\frac{TC}{3} + (C-T)\,p}$$

Of course, the Mises criterion is unaffected by pressure.

The pressure testing data, also from Pai [7.1], are at 1 kBar pressure = 101.4 MPa, which is about twice the compressive strength of the material. The data and predictions for the strengths under pressure are shown in Table 7.1.

It should be noted that the original data were given in psi units to two significant figures. The Mises criterion is calibrated by the tensile strength, as is conventional. The total square errors for the properties predictions compared with the data are given by

Failure criterion 5.5% error
Mises criterion 45.5% error

It is seen that the worst of the three predictions by the Mises criterion is for uniaxial compression, and it differs from the data by a factor of 2.0. Incidentally, the total square error for the predictions from the C–M criterion is 15.4%.

7.2 Brittle Polymers

This example is similar to that of the example in Section 7.1, but for a different materials type that moves a little more toward the brittle end of the scale. Specifically, the materials type is typified by that of PMMA, which

is a transparent, rather brittle thermoplastic, widely used as an inexpensive replacement for glass, and often used in model experiments, especially for dynamic crack propagation. Although as a plastic it is considered as brittle, its strength ratio $T/C \cong 1/2$ places it as intermediate on the overall scale for all materials. The uniaxial and shear failure criterion under pressure are the same as those given in the example in Section 7.1, and the strengths for the C–M criterion under pressure are given by

C–M criterion

$$\sigma_{11}^{T} = T + \left(1 - \frac{T}{C}\right)p$$

$$\sigma_{11}^{C} = -C + \left(1 - \frac{C}{T}\right)p$$

$$S = \frac{TC + (C - T)p}{T + C}$$

It is interesting to note that the C–M forms give a linear dependence on pressure p, while the failure criterion forms in the example in Section 7.1 give a square-root dependence on p. The strengths at $T/C = 1/2$ and for pressure p = 2C (which is of about the same pressure ratio in the example in Section 7.1) are given in Table 7.2.

The greatest difference between these is for uniaxial compressive strength at a factor of

$$\eta = 1.40$$

Thus even at a reasonable pressure of $p = 2C$ these polymers show a wide difference between the two predictions. The differences widen

Table 7.2 Brittle polymer strengths under pressure

	$\frac{\sigma_{11}^{T}}{C}$	$\frac{\sigma_{11}^{C}}{C}$	$\frac{S}{C}$
Failure criterion	1.64	–2.14	1.08
C–M criterion	3/2	–3	1

even further at larger pressures for all three strengths. Furthermore, the effect of the pressure has a very strengthening effect on the strengths themselves.

7.3 Glasses

The problem is that of pushing a curved indenter into an elastic but brittle material such as a glass or ceramic. This is the classical Hertz contact problem.

If the material were a ductile metal, the yielding would begin straight under the indenter, but at some depth. The interest here is with glass, and finding the location of its initial failure, and finding the load at which failure occurs. Take the indenter to be rigid and spherical of radius R (Fig. 7.1). It is found that for the properties (Poisson's ratio) of the brittle material of the half space, both failure criteria maximize at the edge of the contact zone. At this point the surface tractions vanish, and the in-plane stresses are

$$\sigma_r = \frac{(1-2\nu)}{3} q_0$$

$$\sigma_\theta = -\frac{(1-2\nu)}{3} q_0$$

where q_0 is the maximum pressure immediately under the indenter and on its center line. The present failure criterion is controlled by the fracture form, and the analysis of the problem gives the force P acting on the indenter at failure as

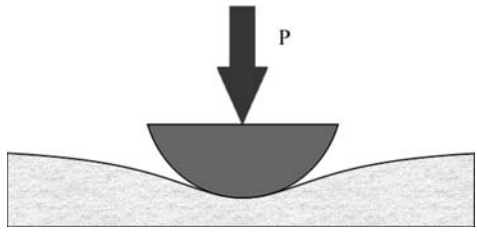


Fig. 7.1 *Spherical indenter on deformable medium.*

Fracture criterion

$$P = \frac{9\pi^3}{2}\frac{(1-\nu^2)^2}{(1-2\nu)^3}\left(\frac{R}{E}\right)^2 T^3$$

C–M criterion

$$P = \frac{9\pi^3}{2}\frac{(1-\nu^2)^2}{(1-2\nu)^3}\left(\frac{R}{E}\right)^2 \left(\frac{T}{1+\frac{T}{C}}\right)^3$$

For $T/C = 1/5$, which corresponds to a high-quality glass such as S-glass, the two predictions of the load at failure differ by a factor of

$$\eta = 1.73$$

Timoshenko and Goodier [7.2] note that many indentation experiments on glass have verified this location for the failure. In fact, it even goes back to the original experiments of Hertz.

7.4 Ceramics

In this example the effect of a two-dimensional pressure upon the shear strength will be examined. The material of application will be for a typical ceramic. Ceramics function poorly under tensile stress states, though they are very useful under combinations of shear stress and compression. Consider a thin layer of a ceramic as a heat shield and subject to pressure in its plane along with shear stress. In contrast to the example in Section 7.2, here both the failure criterion and the fracture criterion will be found to play a role. The two-dimensional pressure $\sigma_{11} = \sigma_{22} = -p_{2D}$, $\sigma_{33} = 0$ will be taken in the same plane as the shear stress σ_{12}, having strength S. The governing criteria are

Failure criterion

$$3S^2 = TC - p_{2D}^2 + 2(C - T)\, p_{2D}$$

Fracture criterion

$$S = T + p_{2D}$$

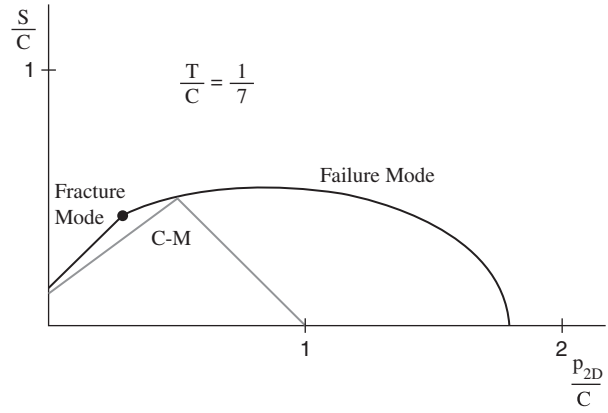


Fig. 7.2 *Shear strength dependence on two-dimensional pressure.*

C–M criterion

$$For \quad p_{2D} \leq \frac{C}{2}$$

$$S = \frac{TC + (C - T)\, p_{2D}}{T + C}$$

$$For \quad p_{2D} \geq \frac{C}{2}$$

$$S = C - p_{2D}$$

At $T/C = 1/7$, the failure envelopes are as shown in Fig. 7.2.

The shortcoming of the C–M criterion in part relates to the fact that it does not account for the intermediate principal stress, σ_2, in its formulation.

7.5 Minerals

This problem is motivated by drilling through hard rock in exploring for water or hydrocarbons. The drill bit causes an intense local state of stress that is primarily composed of shear stress under pressure. The difference between this example and the example in Section 7.4 is that here the pressure is three-dimensional, p, rather than two-dimensional pressure. The governing failure criteria are

Failure criterion

$$S = \sqrt{\frac{TC}{3} + (C - T)\, p}$$

Fracture criterion

$$S = T + p$$

C–M criterion

$$S = \frac{TC + (C - T)\, p}{T + C}$$

The plot of S versus p is given in Fig. 7.3 at $T/C = 1/15$.

There is seen to be a major difference in behavior for three-dimensional pressure from that which occurred in the example in Section 7.4 with two-dimensional pressure. In this example the C–M criterion is mostly unconservative relative to the present failure criterion, whereas in the example in Section 7.4 it is conservative. The present example is much

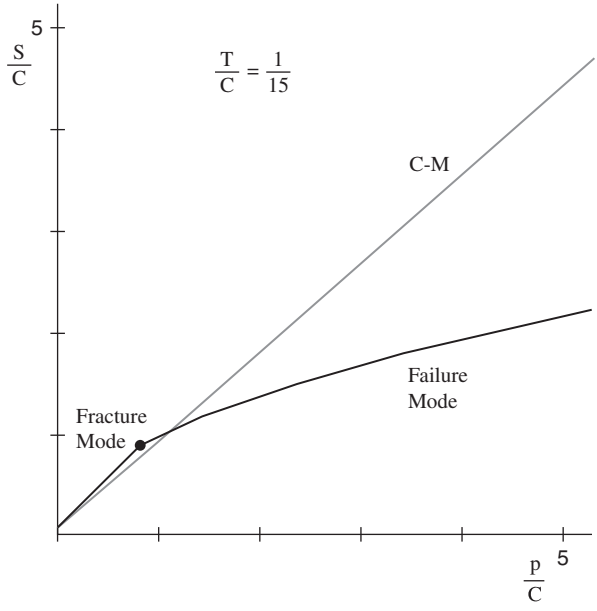


Fig. 7.3 *Shear strength dependence on three-dimensional pressure.*

like that in Section 6.3, where the test data for dolomite were found to bifurcate away from the fracture mode and follow the failure mode as pressure increases, rather than following the straight line projection of Coulomb–Mohr.

7.6 Geo-Materials

This is the problem of a tunnel buried under a large overburden, or equivalently a cylindrical cavity under far-field hydrostatic pressure, Fig. 7.4.

The tunnel is of radius a, and the failure is found to always occur on that surface. The fracture criterion does not enter into this problem, and the other criteria are given by

Failure criterion

$$p = \frac{1}{2}\left[C - T + \sqrt{T^2 - \frac{2}{3}TC + C^2}\right]$$

C–M criterion

$$p = \frac{C}{2} \text{ (independent of } T)$$

At a geo-material value of $T/C = 1/25$ the failure criterion gives

$$p = 0.974C$$

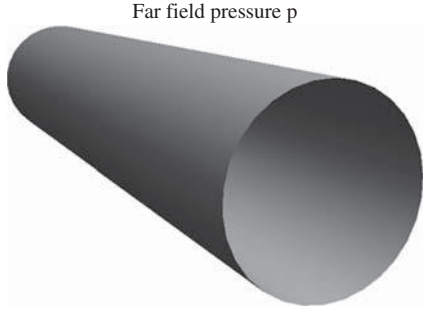


Fig. 7.4 *Tunnel acted upon by overburden.*

At this T/C the two criteria differ by a factor of

$$\eta = 1.95$$

From the solution for the failure criterion it appears superficially that the tunnel can support a greater pressure at $T/C = 0$ than it can at $T/C = 1$:

$$p = \frac{C}{2} \quad at \quad \frac{T}{C} = 1$$

$$p = C \quad at \quad \frac{T}{C} = 0$$

This non-dimensionalized behavior is not only reasonable, it is necessary. The mean normal stress has no effect on the strength at $T/C = 1$, but for lesser values of T/C a compressive mean normal stress does have a strengthening effect. Also, it must be remembered that as the material T/C diminishes, the absolute values for T and C individually also diminish greatly due to the greater degree of damage that exists at smaller T/Cs over the spectrum of materials.

As seen from the examples, sometimes the Coulomb–Mohr criterion is conservative and sometimes unconservative relative to the theory of this book and its anticedants. It is not possible to state a general rule that covers all cases. However, when all the principal stresses are predominately positive or all predominately negative, then it can be shown that the Coulomb–Mohr criterion is always unconservative—sometimes extremely so—as in the example in Section 7.5. In contrast, the present failure theory relies upon its physical development with experimental verification for realistic applications.

Future world-wide applications will be designed to integrate the basic function with life-cycle costs, efficiency, safety, durability, and possibly several other requirements. Specific applications that reflect the new approach will probably include green aircraft, ultra-efficient automobiles, durable wind-turbines, and many other emerging technologies. Such applications as these will in many cases necessitate the development and qualification of entirely new materials. The opportunity to understand and apply materials failure theory to these problems is completely open, and the incentive is almost unlimited.

Problem Areas for Study

1. A common approach for the use of ceramic materials is simply to avoid all applications involving tensile-stress components. How can a more rational approach be devised through using failure criteria but as specialized to ceramics and other similar materials?
2. The present treatment of failure criteria as well as most treatments are of a time-independent nature. How should elasticity (time-independent) type failure criteria be modified to incorporate time-dependent, viscoelastic effects? How does creep rupture fit in with these categories?
3. One of the likely future opportunities is in the area of very large-scale computations and information technology, IT. This is quite different from posing a single, idealized, and isolated boundary value problem and then seeking the related unique limit of strength performance. Probably strength characterization will be used in ways not yet imagined, but it is almost certain that IT will contain strength subroutines in some form or other. Can this be based on specific information rather than just on anecdotal evidence without losing the essence of the IT approach?
4. Although the Coulomb–Mohr theory has no credibility for cohesive engineering materials, it remains as a useful form for soil mechanics and some granular materials. It is the most credible two-property (parameter) form for these latter types of flows. Is there any relationship between soil mechanics and the cohesive engineering materials classes, or are they irretrievably separate and distinct?
5. The failure criteria derived here are implicitly intended for application to materials that either fail in a brittle manner or show strain hardening. Is there any reason why these failure criteria cannot also be applied to materials that show strain softening, providing the failure state is suitably identified?

References

[7.1] Pae, K. D. (1977). "The Macroscopic Yielding Behaviour of Polymers in Multiaxial Stress Fields," *J. Mats. Sci.*, **12**, 1209–14.

[7.2] Timoshenko, S. P. and Goodier, J. N. (1970). *Theory of Elasticity*, 3rd ed., McGraw-Hill, New York.

8

The Ductile/Brittle Transition for Isotropic Materials

In setting out to theoretically characterize and differentiate ductile behavior from brittle behavior it is necessary to have first a credible theory of failure as the starting platform. Historically there was one exceptional mathematical failure criterion and a long list of failures. Perhaps the failures would be better described as misguided attempts. In any case, none of them were successful. Only the Mises criterion (and to a much lesser extent the Tresca criterion) for (but only for) ductile metals was an unqualified and very important contribution.

All other materials types such as polymers, ceramics, glasses, brittle metals, and minerals did not yield to successful mathematical failure criteria descriptions. The cases of non-metallic materials have usually been approached with empirical failure criteria on a case-by-case basis. Such empirical forms do not have established limits of validity, and represent little more than curve fits over narrow ranges of data.

For a general approach to defining ductile-versus-brittle behavior to be successful it must have a rational basis. More specifically, the ductile and brittle material descriptors must be closely interwoven with a general treatment of failure criteria that applies across all materials classes. All of these matters will be taken up now.

8.1 Introduction

The logical place to begin is with an explicit failure criterion that has proven generality. The present failure criterion possesses the requisite generality. It is for homogeneous and isotropic materials, and is composed of two separate criteria. The first part covers the full range of T/Cs, while the second part is fracture-related and covers a partial range of T/Cs.

Polynomial-invariants failure criterion

$$\left(1-\frac{T}{C}\right)(\hat{\sigma}_1+\hat{\sigma}_2+\hat{\sigma}_3)+\frac{1}{2}\left[(\hat{\sigma}_1-\hat{\sigma}_2)^2+(\hat{\sigma}_2-\hat{\sigma}_3)^2+(\hat{\sigma}_3-\hat{\sigma}_1)^2\right]\leq\frac{T}{C} \tag{8.1}$$

and

Coordinated fracture criterion

$$\left.\begin{aligned}\hat{\sigma}_1&\leq\frac{T}{C}\\ \hat{\sigma}_2&\leq\frac{T}{C}\\ \hat{\sigma}_3&\leq\frac{T}{C}\end{aligned}\right\}\ \textit{Also applicable if}\ \frac{T}{C}\leq\frac{1}{2} \tag{8.2}$$

where principal stresses are employed, non-dimensionalized by C.

It could be tempting to assign failure criterion (8.1) as being ductile and (8.2) as being brittle. That, however, would be much too superficial. The problem is more complex than simply declaring a given failure criterion or parts thereof as being either ductile or brittle, as is easily demonstrated.

The individual values of T and of C tell only the specific strength properties for a particular material, but it is the T/C ratio that signifies the materials type that is so pivotal in formulating and in interpreting this general theory of failure. A little later the T/C spectrum will also be found to be of first importance in ductile/brittle matters.

The two coordinated failure criteria for isotropic materials, (8.1) and (8.2), will now be used to develop a framework for characterizing and quantitatively differentiating ductile behavior from brittle behavior. Before doing so, a further brief note on the historical search for failure criteria may be in order.

Criteria (8.1) and (8.2) reduce to the Mises criterion at $T/C=1$, and they also reduce to a physically meaningful brittle limit at $T/C=0$. Thus the Mises criterion is a functional part of this general form. The maximum normal stress criterion, originally proposed by Lamé and Rankine, has a superficial resemblance to the fracture criterion (8.2). If (8.2) were to be considered as a free-standing failure criterion, it would be necessary to omit (ignore) the condition for its applicability shown in (8.2).

In some current as well as past sources of information it is stated that the Mises criterion covers ductile materials and the maximum normal stress criterion covers brittle materials. The first half of this assertion is partially correct, but the second half of it is completely fallacious. For example, common iron is conventionally considered to be a brittle material. When one compares the maximum normal stress criterion with typical data for the failure of iron, the lack of correlation is glaringly obvious. The comparison of the coordinated failure criteria (8.1) and (8.2) with brittle material behavior is quite good; see Section 6.3.

With regard to specific criteria for the ductile/brittle transition, the research directions have related mainly to the activation of dislocations at crack tips in crystalline materials, Rice and Thomson [8.1], Hirsch and Roberts [8.2], and many other contributions of this type. While this direction is very interesting and helpful, it is materials type specific and does not relate to general criteria that apply to all materials types, as are of interest here. Materials science books usually include interesting discussions on ductile and brittle behaviors, but they are almost always in the most general terms, with no specifics on the ductile/brittle transition.

As a final general observation, it is indeed surprising that the association of the T/C strengths ratio with the controlling physical behavior and with the failure types and materials types was not probed or at least recognized a great many years ago. But that is just a curiosity—a footnote to history. Building up a viable theory of behavior is the only thing that really matters, whenever it is accomplished.

8.2 Conventional Difficulty with Characterizing Ductility

The usual methods for characterizing ductility and brittleness are as follows. Brittleness is usually taken to be the sudden, unexpected interruption of the linear elastic stress–strain loading form. Not only is this "unannounced" failure occurrence annoying, it often is dangerous—sometimes extremely dangerous. Some warning of impending failure would be tremendously helpful. This simple classification of idealized brittle behavior will suffice for initial purposes.

Ductility is much more complex. The usual method of identifying ductility is with imposed uniaxial tension. If the material proceeds with a strain hardening state after exceeding some yield stress, and if it is

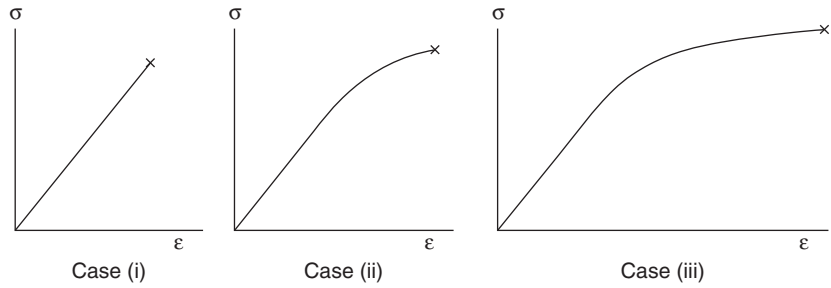


Fig. 8.1 *Failure types.*

capable of experiencing large strains in that state, then it is said to be ductile. This view is unduly restrictive, and certainly does not suffice for describing ductility for three-dimensional stress states. There is a need for a more inclusive definition of ductility.

For present purposes, relevant examples of brittle and ductile behaviors are shown in Fig. 8.1. In these diagrams, σ and ε are taken to represent any stress–strain loading state, one-, two-, or three-dimensional. But the progressive loading is always taken to be of the proportional type. That Case (i) is brittle is obvious. So too is Case (iii) obviously ductile. However, Case (ii) is where the complexity arises. Case (ii) is crucial, and will be first approached in the manner now shown.

The distinguishing characteristic of ductility is taken here to be the capacity to absorb mechanical energy and to convert it into irrecoverable heat energy, or involve some other irreversible damage process. Perhaps the best way to characterize ductility is that it involves mechanical energy dissipation, which is completely absent in the case of idealized brittle failure.

Obviously, Case (ii) has a smaller degree of ductility than does Case (iii). How important is this? Any ductility is much more than no ductility. Even in this sense a small degree of ductility is very helpful, and it affords an important capacity for damage tolerance. In non-uniform stress states, such as stress concentrations, some ductility allows material mobility to adjust to local conditions without causing macroscopic, catastrophic failure.

With the above considerations in mind, Case (ii) will be taken with the following method of broadly (not precisely) distinguishing ductile from brittle behavior. Fig. 8.2 shows Case (ii) with some designations.

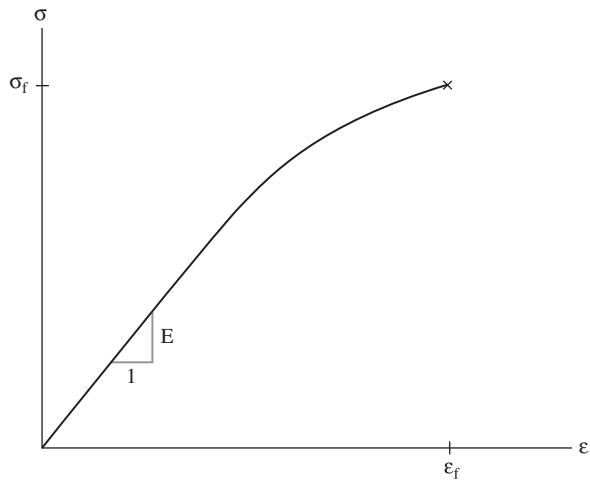


Fig. 8.2 *Failure designations.*

The conditions of brittle-versus-ductile behavior are taken to satisfy

$$\varepsilon_f \cong \frac{\sigma^f}{E}\,, \quad \textit{Brittle}$$

$$\varepsilon_f \gg \frac{\sigma^f}{E}\,, \quad \textit{Ductile}$$

These forms give a general and conventional threshold for ductile behavior. In the second of these the strain at failure may be as little as only 30%, 40%, or 50% greater than the elastic projection for there to be the inception of significant ductility. Otherwise, the mechanical failure behavior is taken to be brittle as in Case (i), or close to it. So, brittle behavior allows a moderate generalization beyond that shown in Case (i). Still, there is a "grey" area, an uncertain area, between the behaviors of totally brittle and ideally ductile types. Also, this operational definition of, and distinction between, ductile and brittle behaviors is stress–state dependent, and is taken to apply to all stress states.

Again, it should be emphasized that even a small degree of ductility can be extremely advantageous, especially as compared with idealized brittle behavior, Case (i). These designations will suffice for having a physical understanding of the possible and probable differences between ductile and brittle behaviors.

The clear message is that ductile-versus-brittle failure behavior is not just an on/off switch. There is a graded and graduated change from one to the other. To say that these are difficult matters would be an understatement. They pose longstanding, classically difficult, unresolved issues.

The fundamental task is that of quantifying ductility levels. None of this so far accomplishes that purpose. To make progress on theoretically gauging ductility levels requires a drastic change in direction and approach from the vague and rather loose traditional discussions of ductility at the macroscopic level. Now a new direction will be sought. In so doing, it should be recognized that any real improvement in characterizing ductility (or lack of) for homogeneous materials must be applicable to all stress states, not just carefully selected, particular ones.

8.3 The Ductile/Brittle Transition

The long-established general consensus is that a given material is taken to be ductile or brittle, based upon experience and exposure. In this heuristic but eminently practical approach, ceramics are said to be brittle, some metals are brittle and some are ductile, and polymers are "all over the place." There is nothing inherently wrong with intuitive and practical approaches, but it is likely that there could be a more organized approach to this problem. First there should be some measurable but non-dimensional physical characteristic(s) that determine the materials type as it relates to its expected failure type, ductile or brittle. Not surprisingly, this missing identifier will be taken here to be the strength ratio T/C.

But if one tries to use only the T/C tag to make general assertions about ductile-versus-brittle failure, one must also ask if this could be true under all conditions. It is quickly seen that there must be qualifiers—especially that of the stress state under which the failure characteristic is to be determined. Generally tensile-type stress states are far more likely to promote brittle behavior, for a given material, than are generally compressive-type stress states. The objective here is to find a rational method for classifying and quantifying expected failure behavior, ductile or brittle with appropriate gradations between, based upon the materials type, specified by T/C, and the specific imposed state of stress at failure.

It will be advantageous to proceed initially in the two-dimensional subspace of the general three-dimensional principal stress space. This will allow ductile-versus-brittle divisions to be identified at the corners resulting from the intersection of two different modes of failure. This would not be possible in three dimensions, because that case has too many dimensions (over-dimensioned) to allow such identifications. After establishing the ductile/brittle division in two dimensions it will be found to be easily generalized to three dimensions.

Take the two-dimensional state of biaxial failure stresses as shown in Fig. 8.3. The two failure criteria (8.1) and (8.2) at $T/C = 1/3$ give the envelopes shown.

Relations (8.1) and (8.2) produce the intersections of failure modes displayed in Fig. 8.3. The domain is divided into ductile and brittle regions at the intersection of the failure modes. If the ductile brittle division were taken through the other intersection corners nearer the origin in Fig. 8.3 it would quickly be found to lead to unacceptable results. For values of T/C other than that in Fig. 8.3 the division between ductile and brittle regions does not pass through the origin as it does in Fig. 8.3. This process is repeated, taking T/C as a continuous variable to find the functional

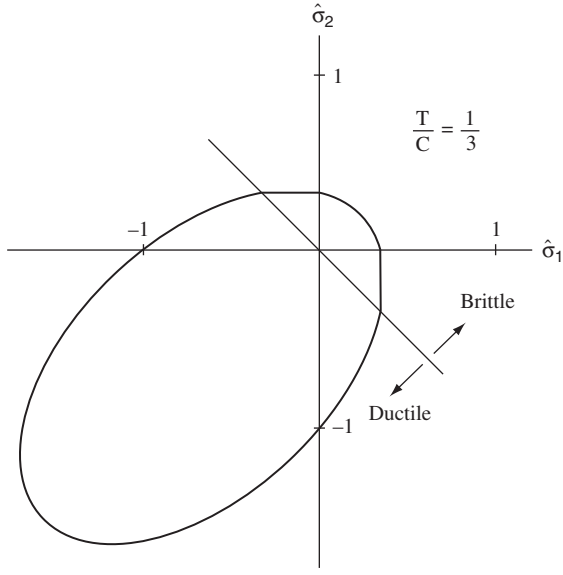


Fig. 8.3 *Ductile/brittle transition, two dimensional example.*

relationship between T/C and the division into ductile and brittle regions. The functional relationship is then taken to have analytic continuation into the range where $T/C \geq 1/2$. The end result is the ductile/brittle criterion to be presented next.

The algebraic details for the process described above for the biaxial stress states, such as is shown in Fig. 8.3, are given in Christensen [8.3]. The resulting general criterion for ductile-versus-brittle failure is found to be

$$\begin{aligned} I_1 &< 3\,T - C, \quad \textit{Ductile} \\ I_1 &> 3\,T - C, \quad \textit{Brittle} \end{aligned} \tag{8.3}$$

where $I_1 = \sigma^f_{11} + \sigma^f_{22} + \sigma^f_{33}$ is the first invariant in terms of the failure stresses. The two-dimensional derivation did not have the third stress component in the first invariant. After obtaining the two-dimensional results, the first invariant was expanded to the three-dimensional form stated above. These relations, for specified values of T and C, then designate which stress states are expected to represent brittle failure and which ones have ductile failure. The first invariant is at its value on the failure envelopes (8.1) or (8.2), and is thus expressed in terms of T and C for the imposed stress state.

Now express relations (8.3) in non-dimensional form for further use here,

$$\begin{aligned} \hat{\sigma}^f_{ii} &< 3\frac{T}{C} - 1, \quad \textit{Ductile} \\ \hat{\sigma}^f_{ii} &> 3\frac{T}{C} - 1, \quad \textit{Brittle} \end{aligned} \tag{8.4}$$

where $\hat{\sigma}^f_{ii}$ is the non-dimensionalized first invariant of the failure stress state from (8.1) or (8.2). The stresses are non-dimensionalized by dividing by C. The relations in (8.4) actually only give the transition from ductile to brittle conditions when the inequalities are replaced by an equality. Thus, for a particular stress state the value of T/C can be found at which the stress state is at the borderline between ductile and brittle failures, but it tells nothing about the relative scales of ductility or brittleness in general failure conditions. This matter will be taken up in the next section.

It is important to visualize the ductile/brittle criteria (8.4) in three-dimensional, principal stress space. The failure criterion (8.1) takes the form of a paraboloid in principal stress space. The fracture criterion, for $T/C \leq 1/2$, cuts off "slices" from the paraboloid, leaving three flattened surfaces on it. The ductile/brittle criteria (8.4) at the transition represents a plane which is normal to the axis of symmetry of the paraboloid and cuts across the flattened surfaces resulting from (8.2). The ductile/brittle transition plane from (8.4) is explicitly given by

$$\hat{\sigma}_{ii}^{f} = 3\frac{T}{C} - 1 \ , \quad D/B \ \ \textit{Transition} \tag{8.5}$$

It is very easy to use (8.5) to determine the value of T/C at the ductile/brittle transition for any given stress state at failure. This will next be illustrated for uniaxial tension and compression. A completely independent derivation of (8.5) is outlined in Problem Area 4 at the end of this chapter.

In uniaxial tension the left-hand side of (8.5) is simply T/C, and when that is substituted into (8.5) the ductile/brittle transition is found to be at $T/C = 1/2$. A broad variety of materials having values of T/C on either side of 1/2 are found to be in general agreement with the ductile-versus-brittle designations from (8.4). This constitutes a validation of the ductile/brittle transition specified by (8.5). Implicitly, this is also a further corroboration of the failure criteria (8.1) and (8.2). It is a useful learning experience to assemble the uniaxial testing data for different materials needed to carry out explicitly this comparison and assessment.

For uniaxial compression, the left-hand side of (8.5) is simply –1, and it follows that the ductile/brittle transition is at $T/C = 0$, so that uniaxial compression is always ductile for all values of T/C, for all materials. Just because the case of uniaxial compressive failure remains ductile for all values of $T/C > 0$ does not mean that there would be expected to be extensive plastic flow in compression. It simply means that in ranges where the uniaxial tension is brittle, the corresponding uniaxial compression is much less so. A small amount of plastic flow in compression is still much more than no plastic flow in tension.

As a further simple example of relations (8.4) and (8.5) consider the case of the failure stress in simple shear. Take S as this failure stress.

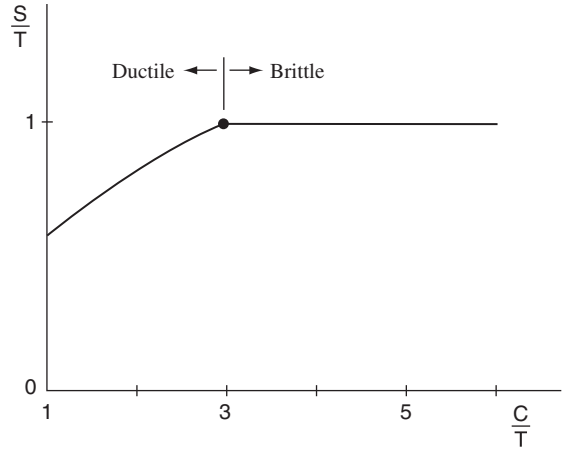


Fig. 8.4 *Shear stress failure types.*

Using relations (8.1), (8.2), and (8.4), along with $S/T = (S/C)(C/T)$ gives the result in Fig. 8.4.

The change in failure mode type as well as the type of failure (ductile or brittle) occurs at $T/C = 1/3$ for a state of simple shear. Other cases will be presented in the next section.

It is important to examine the ductile/brittle characteristic at the brittle limit $T/C = 0$. This is shown in Fig. 8.5.

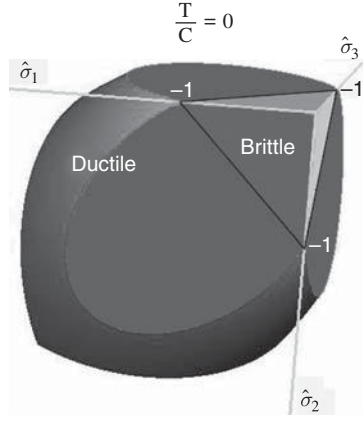


Fig. 8.5 *Ductile/brittle transition at the brittle limit.*

Even in this limit case, $T/C = 0$, there still are separate ductile and brittle regions. At any T/C value the ductile/brittle plane cuts across the three elliptical flattened surfaces (when they exist, $T/C \leq 1/2$) at their 1/4 points nearest the apex of the paraboloid.

At the ductile limit, $T/C = 1$, the ductile/brittle transition has already been displayed in Fig. 5.7.

It would not be expected that the ductile/brittle transition at any T/C value is of a sharp nature. This is not a first-order thermodynamic transition. The significance of this division boundary is that the proclivities toward ductile-versus-brittle failure are in balance, while on either side of the division the weighting gradually shifts to a dominant effect, one way or the other.

8.4 The Failure Number for Gauging Ductility Levels

Planes of equal but unspecified ductility would be parallel to the ductile/brittle transition plane defined by (8.5). It would be advantageous to have the means to characterize the degree of ductility (or lack of it) for particular failure states on either side of the D/B transition plane. To this end, rewrite (8.5) in an alternative form. Take ψ to be specified by

$$\psi = 3\frac{T}{C} - \hat{\sigma}_{ii}^{f} - 1 \tag{8.6}$$

When $\psi = 0$, (8.6) just reverts back to the criterion (8.5) for the transition. But when $\psi \neq 0$, (8.6) has an interpretation as a measure of the degree of ductility or brittleness, as will be shown. This is not surprising, since it is already known that $\psi = 0$ is one point on the ductility/brittleness scale. Positive values of ψ reflect ductility, while negative values indicate brittleness. For a given value of T/C and a given stress state at failure, ψ represents a measure of distance from the ductile/brittle transition plane to the parallel plane containing the stress state of interest.

Examine the case of uniaxial tension. For this stress state $\hat{\sigma}_{ii}^{f} = T/C$, and so (8.6) gives

$$\psi = 2\frac{T}{C} - 1 \tag{8.7}$$

For the range on T/C as $0 \leq T/C \leq 1$ then

$$-1 \leq \psi \leq 1 \tag{8.8}$$

Also, it should be noted that for uniaxial tension the ductile/brittle transition is at $T/C = 1/2$. The stress state of uniaxial tension is the only one that has these matching symmetry forms on T/C and on ψ. Uniaxial tension will be taken as the baseline stress state, and all other stress states will be calibrated relative to it. It comprises the basic template for characterizing ductility. The value $\psi = 1$ corresponds to full, perfect ductility, while $\psi = -1$ corresponds to total brittleness.

Now consider the state of simple shear stress. In this case $\hat{\sigma}^f_{ii} = 0$ and (8.6) becomes

$$\psi = 3\frac{T}{C} - 1 \tag{8.9}$$

For uniaxial tension the ductile/brittle transition value of $T/C = 1/2$ gives $\psi = 0$. But for simple shear the same $T/C = 1/2$ value gives $\psi = 1/2$. This means that at this value of T/C, simple shear failure is more ductile than is uniaxial tensile failure. And at $T/C = 2/3$, then for simple shear $\psi = 1$, the fully ductile situation. Thus for the case of simple shear, values of T/C larger that 2/3 simply remain as fully, perfectly ductile at $\psi = 1$.

Any stress state that at a particular value of T/C gives a value of ψ larger than 1 will be assigned as fully ductile, at $\psi = 1$. Correspondingly, any stress state with a value of ψ less than –1 will be taken as totally brittle at $\psi = -1$. In effect, the planes of perfect ductility and total brittleness are located at fixed distances from the D/B plane for all T/Cs.

It is convenient to rescale the quantity ψ. For $-1 \leq \Psi \leq 1$, change the variable so that the new scale goes from 0 to 1. To this end, take

$$Fn = \frac{1}{2}(\psi + 1) \tag{8.10}$$

or using (8.6)

$$Fn = \frac{1}{2}\left(3\frac{T}{C} - \hat{\sigma}^f_{ii}\right) \tag{8.11}$$

Consistent with taking ψ at limits of –1 and 1, as totally brittle or perfectly ductile respectively, the new quantity Fn will revert to the limits of 0 or 1 when its value from (8.11) is outside these limits. So always

$$0 \leq Fn \leq 1 \tag{8.12}$$

Calibrating to the uniaxial tension stress state is the only one that then gives $Fn = 1/2$ at the T/C values for the ductile/brittle transitions of all the stress states.

A slightly different form of (8.11) is that of

$$Fn = \frac{3}{2}\left(\frac{T}{C} - \hat{\sigma}_m^f\right) \tag{8.13}$$

where $\hat{\sigma}_m^f$ is the non-dimensionalized mean normal stress at failure.

The quantity Fn will be called the failure number. It represents a measure of ductility, with $Fn = 0$ being no ductility, total brittleness, and $Fn = 1$ being full, perfect ductility. Rather than expressing Fn within the 0 to 1 interval, it will always be stated as the corresponding percentage, from 0% to 100%. For $Fn = 50\%$ the material is at the ductile/brittle transition. This is the 50–50 case as regards ductility versus brittleness. For Fn below 50%, brittleness begins to dominate, while for Fn above 50%, ductility begins to dominate.

The procedure for using the failure number is as follows. The sum of the three normal stresses at failure, $\hat{\sigma}_{ii}^f$, are found from the failure theory (8.1) and (8.2), or they may be taken directly from experimental data. These are then put into the failure number formula, (8.11). The two simplest examples are uniaxial tension and compression. The former has $\hat{\sigma}_{ii}^f = \mathrm{T/C}$, and the latter has $\hat{\sigma}_{ii}^f = -1$. Putting these into (8.11) gives the corresponding failure numbers as functions of T/C. These can then be evaluated at specific values of T/C to ascertain the related ductility level.

Table 8.1 shows the values of T/C for the ductile/brittle transition and the failure number formulas for some basic stress states. There are no failures in eqi-triaxial compression in this failure theory.

Next the failure numbers for various stress states are arranged in descending order, thus going from the most ductile stress state to the least ductile, most brittle stress state. Cases over the full range of T/Cs are presented in Table 8.2.

Table 8.1 Failure numbers, Fn

Stress state	D/B transition T/C	Failure number Fn
Uniaxial tension	1/2	$\frac{T}{C}$
Uniaxial compression	0	$\frac{1}{2}\left(3\frac{T}{C}+1\right)$
Simple shear	1/3	$\frac{3}{2}\left(\frac{T}{C}\right)$
Eqi-biaxial tension	1	$1+\frac{T}{2C}-\sqrt{1-\frac{T}{C}+\left(\frac{T}{C}\right)^2}$
Eqi-biaxial compression	Always ductile	$1+\frac{T}{2C}+\sqrt{1-\frac{T}{C}+\left(\frac{T}{C}\right)^2}$
Eqi-triaxial tension	Always brittle	$\frac{\frac{T}{C}\left[1-\frac{3}{2}\left(\frac{T}{C}\right)\right]}{1-\frac{T}{C}}$

Table 8.2 gives a comprehensive view of the ductility levels of the failure mechanisms as a function of the stress state and the materials type. It is the state of uniaxial tension, not shear, that anchors the results. The state of uniaxial tension is of fundamental and singular importance in calibrating the ductile/brittle behavior for all isotropic materials.

It is seen that insofar as ductile-versus-brittle failure behavior is concerned, the sensitivity to the stress-state type is just as great as the sensitivity to the materials type. It must be recognized that there are very large differences in failure behavior for materials below the 50% ductility level from those above it. For example, for the uniaxial tensile stress state, the ductility level of 25% would be in the range of T/Cs for ceramics, while the 75% level would be representative of T/Cs for very ductile polymers. The large majority of cases in Table 8.2 are at or below the ductility level of 50%, indicating problems with expected brittle behavior.

It also is apparent in Table 8.2 that cases below and to the left are dominated by brittle behavior, while those above and to the right are

Table 8.2 Ductility levels Fn, D/B transitions at 50%

Stress state	Ductility levels, Fn				
σ_1:σ_2:σ_3	$T/C = 0$	$T/C = 1/3$	$T/C = 1/2$	$T/C = 2/3$	$T/C = 1$
–1:0:0 Uniaxial compression	50%	100%	100%	100%	100%
1:–1:0 Simple shear	0	50%	75%	100%	100%
1:0:0 Uniaxial tension	0	33%	50%	67%	100%
1:1/2:0 Biaxial tension	0	29%	41%	50%	63%
1:1:0 Eqi-biaxial tension	0	28%	38%	45%	50%

ductile. It is thus possible to identify ductile-versus-brittle failure mode groupings as a function of materials type and stress state. The key to this division are the cases at the 50% level, the ductile/brittle transition. All of the 50% values are exactly at $Fn = 1/2$.

In the four cases in Table 8.2 at $T/C = 0$ with no ductility, total brittleness, it should be remembered that the corresponding stress levels allowed by the failure criteria are at zero, so the ductility levels are irrelevant. However, this does show that exceedingly small T/Cs would have virtually total brittleness. The one case at $T/C = 0$ with a specified ductility level reflects the fact that even at $T/C = 0$ a compressive mean normal stress component does stabilize the material and give an allowable failure stress level with a corresponding ductility level.

It is generally true that for a given stress state, the ductility increases with increasing values of T/C. However, such simple rules become invalid when comparing different stress states. For example, from Table 8.2 it is seen that a $T/C = 1/3$ material in simple shear is more ductile than a $T/C = 2/3$ material in eqi-biaxial tension.

Another set of examples involves more complicated stress states than those given in the above tables. Consider two possible stress states where the principal stresses at failure are in the proportions

$$1 : \frac{2}{3} : -\frac{2}{3}$$

and

$$1 : \frac{2}{3} : \frac{1}{3}$$

It is of interest to know which one is the more ductile of the two stress states, and by how much. To proceed further it is necessary to specify the material of interest. Take the material as a typical, fairly ductile epoxy resin with

$$\frac{T}{C} = \frac{3}{5} = 0.6$$

As the first step, note from (8.11) that the failure number for this epoxy in uniaxial tension is

$$Fn = 60\%, \quad \textit{Uniaxial tension, epoxy}$$

It will be of relevance to not only compare the failure numbers for the two stress states with each other, but also with that for the material in uniaxial tension, which comprises the base line. The answers to these questions of relative ductility levels are certainly not obvious. All three stress states will be shown in the comparison. Using the *Fn* formula (8.11) and the failure criteria (8.1) and (8.2), the respective *Fn* values are as shown in Table 8.3.

It is seen from Table 8.3 that Case I is on the ductile side of the scale, and Case II is on the brittle side. Judging from the values of the *Fn*s, the differences in the degrees of ductility are huge. Furthermore, Case I, the ductile case, is even more ductile than the epoxy itsself is in uniaxial tension. All these results would be pure guesswork without the rationale of the failure number, *Fn*.

Table 8.3 Failure numbers for three-dimensional stress states at $T/C = 0.6$

Case	$\sigma_1 : \sigma_2 : \sigma_3$	$\hat{\sigma}_1^f,\ \hat{\sigma}_2^f,\ \hat{\sigma}_3^f$	Fn
Uniaxial tension	1 : 0 : 0	3/5, 0, 0	60%
I	1 : 2/3 : –2/3	3/7, 2/7, –2/7	69%
II	1 : 2/3 : 1/3	3/5, 2/5, 1/5	30%

The previous examples use a rather ductile epoxy as the material type. An example will now be given in the brittle range. Take a glass type material with

$$\frac{T}{C} = \frac{1}{9}$$

Consider states of simple shear at failure as superimposed upon states of hydrostatic pressure. Again using (8.1), (8.2), and (8.11), the non-dimensional shear stress versus the non-dimensional pressure is found to be as in Fig. 8.6.

The failure number values in Fig. 8.6 are shown at a sequence of pressures. With no pressure the glass material is very brittle in shear. As the pressure increases, the material transforms into an essentially ductile material, becoming very ductile at the larger pressures. Without a reliable theory of ductile and brittle failure it would be very difficult to predict how much pressure would be required to produce a specified degree of ductility improvement for a particular material. It also follows that a state of superimposed hydrostatic tension can convert a ductile material into an essentially brittle one.

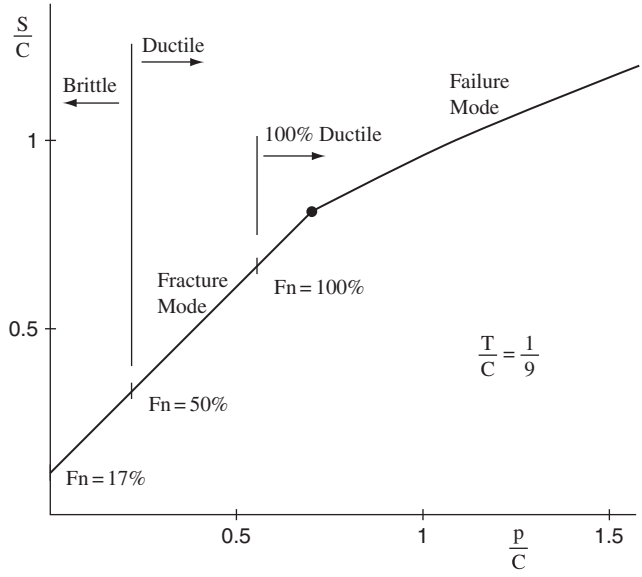


Fig. 8.6 *Shear stress superimposed upon pressure.*

It is also interesting to note that the change of failure mode and the ductile/brittle transition are not coincident, Fig. 8.6. Although such coincidence does occur in two-dimensional stress states, as explained earlier it does not occur with three-dimensional stress states. Thus there is a large region of ductile fracture evident in Fig. 8.6.

The failure number, as specified by (8.11) and by failure criteria (8.1) and (8.2), does give a scale or ranking for the degree of ductility of different stress states for any particular isotropic material specified by its T/C value. This is somewhat like the role and function of the Reynolds number for fluids. Both the Reynolds number, Re, and the failure number, Fn, provide gauges upon the expected behavior due to control/domination by an inherent non-linear physical effect, inertial in the former case and failure in the latter. The difference is that the Reynolds number is for broad classes of flow types, while the failure number is explicit and specific to any particular failure mode.

It would not be expected that the non-dimensional failure number, Fn, would have the same widespread appeal as the Reynolds number does, it is a little too complicated for that. But Fn does reveal that there is an underlying logical basis for predicting the degree of ductility in failure as a function of the materials type and the imposed state of stress. No less than that should be expected from any legitimate and comprehensive failure theory.

Problem Areas for Study

1. Assemble the tensile and compressive uniaxial strength values for a wide variety of isotropic materials. Then form the T/C values and verify that over the range of values from 1 to nearly zero the sequence of values of T/C correlate with the common experience of the ductile-versus-brittle behaviors and the degrees thereof. Reason the same D/B behaviors for the tensile values themselves, and corroborate that for the materials near $T/C = 1/2$ the tensile failure is about at the transition between ductile and brittle behaviors.
2. All ductile/brittle matters are of high relevance and importance. A method of predicting ductile-versus-brittle failure behavior has been given here as related to the materials type (T/C ratio) and the imposed stress state. Is there a completely different macroscopic D/B approach that could be taken to achieve the same objective? Could ductile-versus-brittle behavior be quantitatively related to or correlated with

other physical specifications such as some of the thermomechanical properties?

3. Spallation-type failure due to pressure wave reflection is a common occurrence in impact and detonation problems. Even very ductile materials undergo spallation. Can this be related to ductile-versus-brittle failure behavior as determined by the stress-state type? More specifically, can the ductile/brittle criterion of this chapter be applied to this problem?
4. An alternative derivation of the ductile/brittle transition criterion (8.5) proceeds as follows. A plane normal to the generating axis of the polynomial-invariants paraboloid is taken to divide the failure envelope into ductile and brittle regions. The plane is specified by the first invariant expressed in terms of the failure stresses, as given by the left-hand side of (8.5). It is the location of the plane relative to the origin of the principal stress coordinate axes that will complete the D/B transition specification. Since this failure theory is controlled by only two failure properties, T and C, so too the D/B transition criterion will be taken to be controlled by and only the same two failure properties. This then involves specifying the D/B transition criterion having the first invariant as in (8.5), but the right-hand side of the equation must be given by $[b + d(T/C)]$, where b and d are coefficients to be determined. Thus two independent conditions are needed to find b and d, and they must come from the characteristics of uniaxial tensile and compressive failure. The first condition is that an examination of uniaxial failure data for a wide variety of materials types, as in Problem Area 1 above, shows that the D/B transition in uniaxial tension occurs at $T/C = 1/2$. The second condition must then involve uniaxial compressive failure. This condition is not quite as transparent as that for uniaxial tension, but it also can be reasoned by observational behavior. It is clear that the D/B transition in uniaxial compression must be at a value of T/C that is much less than that for tension. The uniaxial compressive failure at $T/C = 1$ is always of a ductile nature. As T/C is diminished below the value of 1, the common failure observations remain as being ductile, although less so, and possibly approaching less ductile, more nearly brittle behavior for small values of T/C. For this compressive case, there is no clear-cut D/B transition in the way there is for the tensile case. These disparate failure characteristics are reconciled by taking the D/B transition in uniaxial compressive failure as occurring at $T/C = 0$. These two conditions then determine the

coefficients b and d, and it is found that they give exactly the same result as shown in (8.5). In appraising this result, which derivation is preferable or more fundamental: that given in Section 8.3 or that outlined here? Is there yet another independent method of determining the D/B transition criterion (8.5)? Are not these derivations of (8.5) yet another foundational verification of the overall failure theory?

5. The criterion for the ductile/brittle transition is given in Section 8.3, in which I_1 is the first invariant in terms of the failure stresses. In the limit of a perfectly ductile material, $T/C = 1$, the criterion divides the Mises yield cylinder in principal stress space into ductile and brittle regions, as shown in Section 5.3. It follows that a sufficiently large superimposed tensile state is predicted to convert a normally perfectly ductile metal into a brittle material form. This cannot be explained through conventional dislocation theory. What modification to the theory would be needed to explain this? Can the Rice–Thomson theory of the ductile/brittle transition be applied or modified to apply to this problem? Would the explanation necessarily require a dislocation core incorporation?
6. Dislocation theory predicts that resolved shear stress leads to the Tresca criterion. Macroscopic material behavior shows that the Mises criterion is the more proper form. Although there have been explanations for this divergence they are not completely convincing. Is there an all-inclusive explanation, perhaps involving the effects at grain boundaries?
7. A specific criterion for the ductile/brittle transition has been derived in Section 8.3. Does this have any relationship to a second-order transition as is often posed in thermodynamical terms?

References

[8.1] Rice, J. R. and Thomson, R. (1974). "Ductile versus Brittle Behavior of Crystals," *Phil. Mag.*, **29**, 73–97.

[8.2] Hirsch, P. B. and Roberts, S. G. (1997). "Modeling Plastic Zones and the Brittle–Ductile Transition," *Phil. Trans Royal Soc.* (Lond.), **A355**, 1991–2002.

[8.3] Christensen, R. M. (2005). "Exploration of Ductile, Brittle Failure Characteristics through a Two Parameter Yield/Failure Criterion," *Mater. Sci. and Eng.*, **A394**, 417–24.

9

Defining Yield Stress and Failure Stress (Strength)

The concepts of yield stress and strength are extremely widely used and probably equally widely misunderstood. The variations of interpretation for these properties are the norm, not the exception. The treatment of failure in this book employs the uniaxial strengths T and C for calibration. These in turn require some operational definitions in order to extract them from testing data. It is an unfortunate fact of life that almost every individual piece of work that reports yield stress and strength performance levels does so without a clear guide on these matters. Thus there is usually a wide range of property values to choose from, and little advice on how to do it.

The problem is not quite as severe as might at first be surmized from the above preview. For example, whatever method may be used for assigning values to T and C, it will probably not cause undue trouble if it follows that the predictions from the resultant failure theory are interpreted in the same manner and with the same qualifications. Still, that does not answer the basic question. How should yield stress and failure stress be defined in a formal discipline context? This longstanding and long vexing problem will now be addressed.

9.1 Yield Stress and Strength as Historically and Currently Practiced

One of the most obvious but still most difficult problems in dealing with matters of failure is that of defining the yield stress and the failure stress, commonly known as strength. These properties are needed to calibrate failure criteria. Such properties as modulus, yield stress, and strength are widely codified and quoted. Modulus E is straightforward, but the terms "yield stress" and "strength" have a somewhat ambiguous history of use.

In fact, there has not even been agreement on the proper designations for these properties. The terms "yield point," "proportional limit," and "yield strength" have been used for the first property. For the failure stress such terms as "ultimate strength," "strength," "rupture," "limiting stress," and many more have been used. All of these have had rather different interpretations, but none have been entirely satisfactory and none universally adopted.

Modern continuum mechanics does not offer any special insight into these matters, nor does traditional mechanics of materials, which with considerable irony is also called "strength of materials," even though it has almost nothing to say about strength, Timoshenko [9.1]. In materials science, Cottrell [9.2] gives helpful discussions but does not give operational definitions. Hull and Bacon [9.3] state that "the yield stress is not unique" in recognition that the plastic deformation in metals due to dislocation flow is not a singular event but a diffuse process. Similarly for strength, the usual approach is to assign the stress at which the specimen ceases to function as being the strength. The term "ceases to function" usually remains undefined.

Perhaps as the result of these conceptual difficulties, modern usage has evolved into that of an arbitrary rule: the 0.2% strain offset rule for determining the yield stress of metals. For other materials there are not even arbitrary rules; there are only individual preferences and proclivities.

For perfectly brittle materials there is no problem. The yield stress is irrelevant and the strength is obvious. But for partially ductile or very ductile materials there is uncertainty and confusion about how to determine the yield stress and the strength.

It is quite apparent that to make the best use of failure criteria, it is necessary that they be implemented and supported by consistent and meaningful definitions of their calibrating properties and correspondingly of the interpretation of their results. To this end, rational definitions of yield stress and strength will now be sought.

9.2 A Rational Definition of Yield Stress

Consider a typical, ductile material stress strain curve as shown in Fig. 9.1.

The related constitutive form will be taken to be that of the strain-hardening type and applicable to any standard test such as those for uniaxial tension, uniaxial compression, shear, or any proportional loading

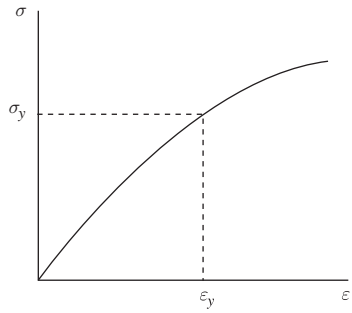


Fig. 9.1 *Stress strain curve.*

state. Now take the first and second derivatives of the stress–strain curve in Fig. 9.1. These are shown schematically in Fig. 9.2.

At the inflection point shown in Fig. 9.2a the second derivative in Fig. 9.2b reaches a maximum.

It is now seen that the strain $\varepsilon = \varepsilon_y$ in Fig. 9.2b, at which the magnitude of the second derivative reaches a maximum, is a transition point, and the associated stress σ_y in Fig. 9.1 is the yield stress. Thus, this definition of the yield stress, σ_y, is the single point on the stress–strain curve at which the tangent modulus, E_T, is changing at the greatest rate with respect to increasing strain. The criterion for the yield stress is then

$$\sigma = \sigma_y \quad at \quad \left|\frac{d^2\sigma}{d\varepsilon^2}\right| = maximum \tag{9.1a}$$

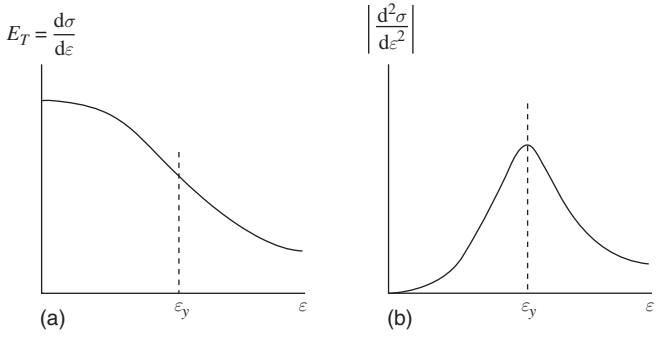


Fig. 9.2 *Derivatives of stress strain curve.*

or

$$\sigma = \sigma_y \quad at \quad \frac{d^3\sigma}{d\varepsilon^3} = 0 \tag{9.1b}$$

This criterion for yield stress was given and examined by Christensen [9.4]. Surprisingly, it apparently was never considered historically.

For ductile metals the location of the maximum of the second derivative is that condition at which the dislocation flow is sufficiently intense and varied as to cause this result. Such a process point involves both dislocation nucleation and the actual dynamics of the flow. For other materials such as glassy polymers, this point of the maximal second derivative is caused correspondingly by molecular rearrangement and damage at both the molecular and macroscopic scales.

The best way of viewing this yield behavior is that the first derivative $\frac{d\sigma}{d\varepsilon}$ versus ε is constantly decreasing for increasing ε, but it has an inflection point that identifies the transition point σ_y. This single unique point then designates the transition from the previous nearly ideally elastic behavior to the following behavior approaching perfectly plastic flow. That the material may never truly attain the perfectly plastic state is usually caused by the intercession of the effects of flow anisotropy or localization, and so on.

To illustrate the process, an analytical form for typical stress strain curves will be taken so that the derivatives in (9.1) can be evaluated. Decomposing strain into elastic and plastic parts, take

$$\varepsilon = \frac{\sigma}{E} - aLn\left[1 - \left(\frac{\sigma}{\sigma_0}\right)^m\right] \tag{9.2}$$

where a, m, and σ_0 are parameters to be specified. It can be shown that as exponent m becomes very large, (9.2) approaches that of the elastic–perfectly plastic case. Thus, except in this limiting case, the form (9.2) represents a continuous function with continuous derivatives in accordance with most physical observations.

Two examples will be given to show the evaluation of the yield stress using (9.1). The two examples with different values for exponent m are

(i)	(ii)
$E = 70$ GPa	$E = 70$ GPa
$\sigma_0 = 500$ MPa	$\sigma_0 = 500$ MPa
$a = 0.01$	$a = 0.01$
$m = 5$	$m = 10$

The corresponding second derivatives from (9.2) are shown in Fig. 9.3.

It is seen that these second derivatives give sharply defined maxima. The actual yield stresses were determined from (9.1b), the third derivative equal to zero.

The full stress strain curves with the yield stresses determined by (9.1) are given in Fig. 9.4.

The yield stresses from (9.1) in the two examples are given by

$$\begin{aligned} m = 5 \quad & \sigma_y = 236\,\text{MPa} \\ m = 10 \quad & \sigma_y = 339\,\text{MPa} \end{aligned}$$

It is apparent from Fig. 9.4 that the yield stress values are larger than the values at which non-linearity can be first observed, but it can also be shown that they are far less than the values from the 0.2% strain offset rule. If the exponent $m = 100$ were taken in these examples, the result would be almost indistinguishable from the elastic–perfectly plastic case.

This new method for determining the yield stress complies with the original objective to find a rational criterion for the operation. However, it has a major disadvantage that renders it as almost unusable. Typical

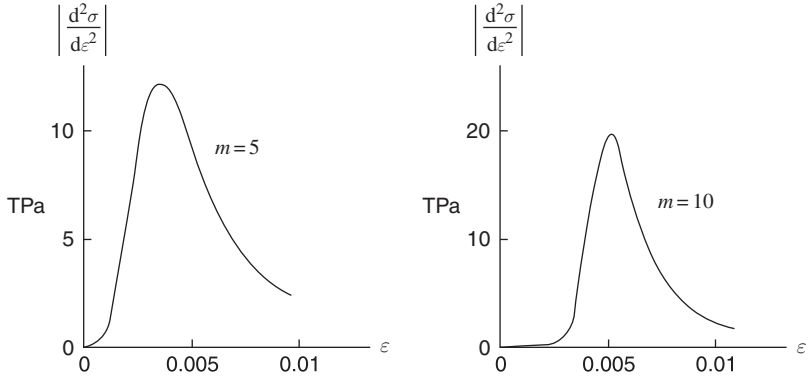


Fig. 9.3 *Second derivatives for the examples.*

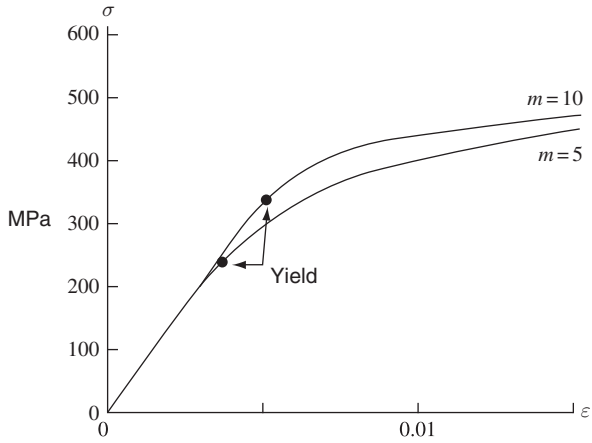


Fig. 9.4 *Yield stresses for the examples.*

stress–strain curves are actually just sequences of datapoints. Reliably determining second and/or third derivatives from such typical datasets is prohibitively difficult. So the situation appears to be a little like the famous words "the operation was a success but the patient may not survive."

The practical means and method for determining the yield stress is still in an unsatisfactory and suspended state. Here is how to proceed further. The condition (9.1) is taken as the rigorous definition of the yield stress, even though it is difficult to implement. An approximation to the criterion (9.1) will be sought in a form that is easy and direct to use, but still retains the essential calibration of (9.1) in the stress–strain problems of interest.

After examining several different forms the following was selected as a simple but reliable operational means of determining the yield stress. The yield stress is taken (designated) as that stress at which the actual strain is 5% greater than that of the linear elastic projection. This definition is

$$\sigma = \sigma_y \quad at \quad \frac{\varepsilon_y - \varepsilon_L}{\varepsilon_y} = 0.05 \tag{9.3}$$

with ε_L being the linear elastic range strain shown in Fig. 9.5.

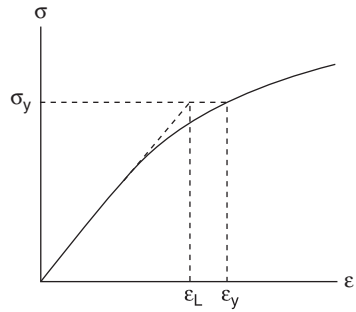


Fig. 9.5 *Strain deviation condition.*

In the two examples considered earlier, the rigorous yield stress given by (9.1) then gives the strain form in (9.3) as

$$For\ m = 5 \qquad \frac{\varepsilon_y - \varepsilon_L}{\varepsilon_y} = 0.066$$

and

$$For\ m = 10 \qquad \frac{\varepsilon_y - \varepsilon_L}{\varepsilon_y} = 0.042$$

These values are close to the 0.05 value of the 5% rule. The 5% strain deviation rule given by (9.3) provides a reasonable and easy-to-use criterion for determining the yield stress. It must be remembered, however, that it is only a convenient approximation to the rigorous definition of yield stress in (9.1). In the examples the approximate yield stresses given directly by the 5% strain deviation rule (9.3) are

$$\begin{aligned} For\ m = 5 \quad & \sigma_y \cong 220 MPa \\ For\ m = 10 \quad & \sigma_y \cong 347 MPa \end{aligned}$$

These approximations compare quite well with the exact values.

9.3 A Rational Definition of Failure Stress

How should one define strength? Is it simply the stress at which a specimen ruptures or fragments into multiple pieces? If the material were very brittle then the resulting failure stress is obvious. When

the material is not perfectly brittle the problem becomes much more complex.

Sometimes at very large strains in uniaxial tension, materials undergo anisotropic reorientation of the very small scale microstructures, and very large stresses can be attained before final rupture. This one-dimensional behavior probably has very little to do with usable strengths in three-dimensional stress states. For example, polymers show strain-induced extreme molecular orientation in one dimension, but it is meaningless for three-dimensional stress states. This complicating but real condition must be dealt with.

First, pose the general situation as shown in Fig. 9.6. From this figure it is surmised that yield stress as defined earlier is almost useless in specifying strength. They are completely independent properties.

To get a grasp on the failure problem, start with the three idealized cases in Fig. 9.7.

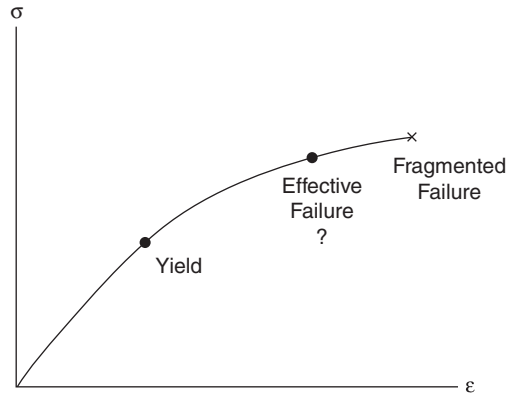


Fig. 9.6 *Failure stress possibilities.*

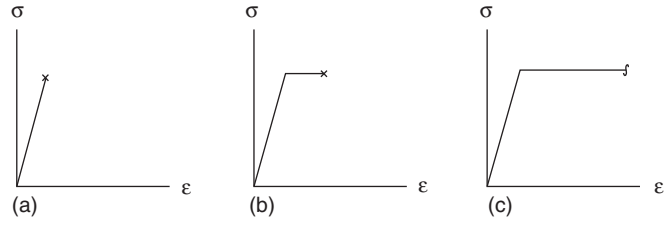


Fig. 9.7 *Idealized stress strain cases.*

Case (a) is the perfectly brittle case, and Case (c) is the perfectly ductile, perfectly-plastic case. What is Case (b)? Is it brittle or ductile or something in between?

Approach the problem by asking, what is the failure strain in this idealized case that represents the inception of full ductility? Ultimately, the answer to this question will help to define strength in the general situation.

Continuing with the idealized cases of Fig. 9.7, take the following designations shown in Fig. 9.8.

U_C *Conserved or recoverable energy*
U_D *Dissipated energy*

The linear partition between the conserved and dissipated energies is that of the idealized unloading line controlled by the slope E.

Take

$$(U_C - U_D)^2$$

as an indicator of the energy dominance or divergence for the two types of energy. Then

$$\begin{aligned} (U_C - U_D)^2 &= U_C^2 \quad at \quad \varepsilon_0 = \frac{\sigma_f}{E} \\ (U_C - U_D)^2 &= 0 \quad at \quad U_C = U_D \\ (U_C - U_D)^2 &= U_D^2 \quad at \quad \varepsilon_0 \gg \frac{\sigma_f}{E} \end{aligned}$$

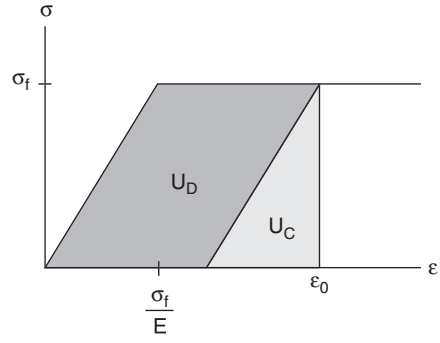


Fig. 9.8 *Idealized cases nomenclature.*

where now ε_0 is taken to be the failure strain. Normalize the above expression by U_C^2 and define the energy dominance by

$$\Lambda = \left(1 - \frac{U_D}{U_C}\right)^2 \tag{9.4}$$

From Fig. 9.8 the two energy terms in (9.4) are given by

$$\begin{aligned} U_c &= \frac{1}{2}\frac{\sigma_f^2}{E} \\ U_D &= \sigma_f \left(\varepsilon_0 - \frac{\sigma_f}{E}\right) \end{aligned} \tag{9.5}$$

Substituting (9.5) into (9.4) gives

$$\Lambda = \left[3 - 2\frac{\varepsilon_0}{\left(\frac{\sigma_f}{E}\right)}\right]^2 \tag{9.6}$$

A graph of (9.6) of the energy dominance Λ as a function of the failure strain ε_0 is shown in Fig. 9.9.

The brittle and ductile designations in Fig. 9.9 are with failure at $\varepsilon = \varepsilon_0$ in Fig. 9.8. The case of $\varepsilon_0 = \frac{\sigma_f}{E}$ in Fig. 9.9 is that of the perfectly brittle behavior. As ε_0 is increased beyond that, greater degrees of ductility

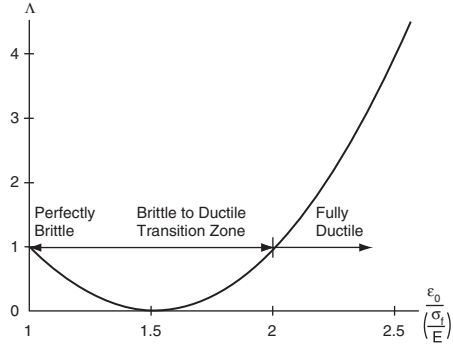


Fig. 9.9 *Energy dominance versus failure strain.*

are imparted to the material. The symmetry characteristic in Fig. 9.9 between

$$\varepsilon_0 = \frac{\sigma_f}{E} \qquad \textit{and} \qquad \varepsilon_0 = 2\frac{\sigma_f}{E}$$

is used to designate the latter value of ε_0 as that at which full ductility is first attained.

$$\varepsilon_0 = 2\frac{\sigma_f}{E} \quad \textit{Inception of full ductility} \tag{9.7}$$

This is in accordance with the ductile/brittle failure behavior and characteristics found in Chapter 8. Beyond $\varepsilon_0 = 2\frac{\sigma_f}{E}$, fully ductile states of failure are assured.

At the inception of fully ductile behavior, as defined above, then from (9.5)

$$\begin{aligned} U_c &= \frac{1}{2}\frac{\sigma_f^2}{E} \\ U_D &= \frac{\sigma_f^2}{E} \end{aligned} \tag{9.8}$$

So at this start of fully ductile behavior there is

$$U_D = 2\,U_c \tag{9.9}$$

The original objective has been accomplished. For the idealized stress strain curve in Fig. 9.8 the failure strain has been found at which a state of full ductility can be designated. Furthermore, the dissipated and the conserved energies are found to satisfy (9.9) at this value of ε_0, giving the inception of full ductility.

Now the final step is taken to move beyond the idealized forms of Figs. 9.7 and 9.8. In actual behavior the effective failure stress is taken at the strain at which the state of full ductility first commences, as given by (9.9) from the idealized case. This general case is as shown in Fig. 9.10.

The effective failure stress in the general case of Fig. 9.10 is defined by

$$\sigma = \sigma_f \quad \text{at} \quad U_D = 2\,U_c \tag{9.10}$$

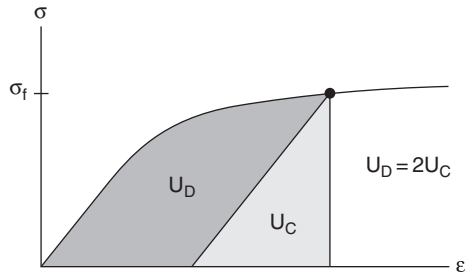


Fig. 9.10 *Strength definition.*

where U_c and U_D are the conserved and dissipated energies shown in Fig. 9.10. If U_D never attains the value $2U_c$ before actual rupture, then the failure stress is of course that at actual rupture. For ductile materials beyond the inception of full ductility, the material begins to develop "texture" which tends to give false indications of usably increasing strength. In this case of ductile materials the failure stress shown in Fig. 9.10 is that which could be called the effective failure stress or strength, since it is not the actual breaking of the specimen but rather a conceptual loss of material function. Although the terms "conserved" and "dissipated" energies are used here in the general case, they are motivated by the more precise definitions of them from the previous idealization in Fig. 9.8.

This definition of effective failure stress may occur before final rupture, or it may actually be that of final failure in the case of brittle behavior. Henceforth, the term "effective failure stress" will be dropped, and simply taken as the strength prescribed by the failure criterion (9.10). In using (9.10) the dissipated energy is that of all non-conservative energy expenditures through damage, dissipation, and any other mechanisms.

The failure criterion (9.10) applies to most materials that possess a linear elastic range of behavior followed by yield and then by failure. This then excludes all elastomers, all polymeric materials above their glass transition temperatures and probably most or all biological materials. It does include all the usual engineering materials, whether the ranges beyond the yield stress are strain-hardening, or brittle, or almost brittle. Strain-softening materials generally require special consideration, though this definition of strength may usefully apply to some cases.

The criterion (9.10) applies to any stress state—tension, compression, shear, or anything else—though it is of the most interest here in

applications to uniaxial tension and compression for use in the failure theory. In this application to uniaxial tension and compression, and within the framework of isotropic failure theory, there is always $T \leq C$. Equality occurs in the perfectly ductile case. Any apparent data that seem to suggest $T > C$ would not be admissible.

9.4 Significance and Conclusions

The yield stress is rigorously defined by (9.1), involving the maximum of the second derivative of the stress strain curve. Taking derivatives from datasets is not an easy to use or reliable procedure. Accordingly a related operational method for approximately determining the yield stress has been found and verified. This method, embodied in (9.3) is that of the 5% strain deviation from linearity. It is very easy to use and far more realistic than the arbitrary 0.2% offset rule usually used with metals.

In contrast to the yield stress, the strength is derived from a drastically different energy-type criterion that results in the definition

$$\sigma = \sigma_f \quad \text{at} \quad U_D = 2\,U_c$$

where U_c and U_D are the conserved and dissipated energies shown in Fig. 9.10. The consequence of this definition is that the effective failure stress for ductile materials exhibiting extensive strain-hardening is not that of the final rupture stress, but rather is a somewhat lesser, intermediate value of stress. Following the reasoning used to derive this failure stress, it is really an effective failure stress—one that is physically transferable to use in other stress states through the failure criterion. This places a wholly new and different interpretation on the definition of the strength. It is also interesting and revealing to note that this strength definition involves the elastic modulus E, or the modulus for whatever stress state is being considered.

The yield stress and the strength definitions are appropriate to engineering materials, all of which have a linear range of elastic behavior. This is followed by irreversible damage or dissipation, ultimately leading to failure. The associated failure criteria are in effect the constitutive relations specifying the termination of the conservative and reversible range of the elastic behavior. Both the yield stress and strength definitions are physically direct, each is quite apparent on the usual computer graphics displays.

The 5% strain deviation rule for yield stress is just a little above the apparent loss of linearity. The failure criterion $U_D = 2U_c$ for strength is also easily calibrated by appearance on the stress–strain curve. Quantification of these follow directly.

The terms "global" and "local" have some meaning when considering these definitions of yield stress and strength. With regard to the full stress–strain curve, the yield stress is a local property dependent upon the null value of the third derivative, as shown in the examples. The strength is a global property involving an integration operation over major portions of the stress–strain curve. When one sees both of these properties in this perspective it is immediately apparent that in most cases the strength property is of far more significance than is the yield stress property. Fortunately, the strength definition involves only an integration operation that is very easy to use in quantifying its value from data.

It is not expected that these new definitions of yield stress and strength will ever come into general use. There is too much momentum embedded in the old intuitive and empirical preferences that are in general use to ever overcome that. It is strongly recommended, however, that in using the failure criteria derived here, they (the failure criteria) should utilize these rational definitions of failure stress or yield stress or other similarly carefully defined ones. Above all, there must be consistent usage of the definitions of yield stress and failure/strength when used as parts of a predictive failure theory.

Finally, it is worth emphasizing further the difference between yield stress and strength. The failure criteria that are derived and used here are mainly intended for use with the specific conditions of failure, rather than with yielding. It may still be possible to use the failure criteria with considerations of yielding so long as both the calibrating properties and the predicted failure (yield) envelopes are consistently defined and interpreted.

Problem Areas for Study

1. How should yield stress and failure stress be defined? Traditional methods are quite arbitrary. Sections 9.2 and 9.3 address this subject, but no doubt other approaches may also be derived. Will there ever be unique and accepted definitions for these properties?

2. In distinction to the definition of strength given in Section 9.3, can stress at a particular value of strain be used to specify the stress at failure? How would there be any account of the many different types of materials in using this method?
3. The present theory of failure is mainly derived and intended for application to strength characterization rather than to yield stress. Are there any precautions to be observed in applying it to yield stress conditions? What considerations must be taken into account?
4. A longstanding problem is that of *in situ* strength properties versus "virgin" material properties. This is particularly important with fiber composites, though it is common to all materials applications. How should failure criteria be modified, if at all, to reflect these effects?
5. Consider the problem of ideal, intrinsic strength properties. Can the intrinsic strength properties, which would exist ideally if there were no damage, defects, or disorder, be predicted or inferred from strength measurements on existing materials? This would necessarily be limited to the macroscopic mechanics of materials formulations. This would not be comprehensive from a multi-scale general materials point of view. Is this a problem?

References

[9.1] Timoshenko, S. P. (1953). *History of Strength of Materials*, McGraw-Hill, New York, Dover (1983).
[9.2] Cottrell, A. H. (1981). *The Mechanical Properties of Matter*, Krieger, Malabar, FL.
[9.3] Hull, D. and Bacon, D. J. (2001). *Introduction to Dislocations*, Butterworth Heinmann, Oxford.
[9.4] Christensen, R. M. (2008). "Observations on the Definition of Yield Stress," *Acta Mech.* **196**, 239–44.

10
Fracture Mechanics

The field of fracture mechanics is highly developed and widely applied. There are many excellent sources of information on the subject, and an independent treatment here is certainly not needed. There does, however, seem to be some confusion about the separate roles of fracture mechanics and failure criteria. This chapter considers some of the issues, and seeks a clarification of the separate and distinct capabilities of the two disciplines.

The terminology failure criteria, as part of failure theory, and as used in this book, is consistent with common usage. In jointly considering failure criteria and fracture mechanics some basic and obvious questions are as follows. Are failure criteria and fracture mechanics, as commonly practiced, completely independent fields? Alternatively, are they largely independent, but with some degree of overlap?

Another question is foremost in considering fracture mechanics and failure criteria. Is one approach more general and inherently superior to the other, as is sometimes claimed, either way? This coordinated examination of the two fields begins with a highly concentrated summary of the remarkable development of the field of fracture mechanics.

10.1 Fracture Mechanics Development

It all originated with Griffith in 1921. He recognized that flaws could induce failure in materials, and posed and solved the idealized problem of a single crack in an infinite two-dimensional, isotropic, elastic medium under transverse load. The famous solution, obtained from the energy balance principle, is given by

$$\sigma = \sqrt{\frac{2E\gamma}{\pi a}} \tag{10.1}$$

where σ is the far-field stress causing the crack to open and grow unstably under plane stress conditions, with a being the half crack length and γ the classical surface energy due to the breakage of bonds in the generation of new crack surface.

In considering applications to glass it was found that the concept and form of (10.1) is correct, and the prediction of the failure load level was reasonable using the classical surface energy γ for glass. However, for other materials, especially ductile metals, their values for γ can greatly underestimate the actual energy required to extend the crack.

Much later in the 1950s Irwin generalized the form of (10.1) by introducing the macroscopic energy release rate, G, as an independent property. Thus for the same central crack problem

$$\sigma = \sqrt{\frac{EG}{\pi a}} \tag{10.2}$$

where

$$G = \frac{\partial U}{\partial A} \tag{10.3}$$

with U being the total potential and A the crack area.

Going even further, Irwin greatly expanded the utility and applicability of the method by introducing the stress intensity factor with

$$\sigma_{ij} = \frac{K_I}{\sqrt{2\pi r}} f_{ij}(\theta) \tag{10.4}$$

where this is the form of the stress field near the linear elastic square root singularity. K_I in the above-mentioned central crack problem is given by

$$K_I = \sqrt{\pi a}\,\sigma \tag{10.5}$$

In more general problems then,

$$K_I = \alpha\sqrt{\pi a}\,\sigma \tag{10.6}$$

for Mode I crack opening conditions. Similar forms follow for the two shear modes, II and III, and for plane strain. The values of α can be found for any problem of interest with characteristic dimension a.

The criterion for crack instability is then given by

$$K_I^2 = EG \tag{10.7a}$$

or simply

$$K_I = K_{Ic} \tag{10.7b}$$

where K_{Ic} can be viewed as a property of the material, the fracture toughness.

The treatment just outlined follows from linear elastic analysis. Griffith provided the conceptual breakthrough and Irwin the follow through with an imaginatively organized, general methodology. This approach depends upon small-scale yielding/damage as existing only in the vicinity of the crack tip, and is referred to as linear elastic fracture mechanics.

The method recognizes that linear elastic singularities do not really exist, but still creatively uses them as the normalizing agent common to all cases of its type. The entire formalism is elegant and self-contained, and it represented a huge step forward in understanding the failure of materials in the presence of stress-magnifying flaws, defects, and cutouts.

In effect, the above treatment represents the first part of the development of the field. The second part represents the approach to be taken when small-scale yielding does not apply—that is, when the material yielding is over a region of about the same size as the crack or even over a larger region. This situation is especially important with ductile metals, and it will be summarized next. General treatments of fracture mechanics can be found in many sources, such as Kanninen and Popelar [10.1], Broberg [10.2], Suo [10.3], and Anderson [10.4].

Continuing now with the case of large-scale yielding, Rice [10.5] recognized that a wholly different and new approach would be required. He showed that the full plasticity problem could be simulated by the more direct non-linear elasticity problem in the cases of proportional loading. Then, with an innovative use of a path-independent line integral, the J integral, he formulated the fracture problem in a completely different and more tractable manner. In the case of small-scale yielding and for linear elasticity the J integral method gives the value of the contour integral as

$$J = \frac{K_I^2}{E} \tag{10.8}$$

This development then opened the door to using this method in the much more complicated non-linear case (and linear as well) usually using power-law forms to represent the non-linear constitutive behavior. This theoretical synthesis by Rice of all the elements of fracture mechanics completely consolidated the field and further enlarged its utility, as shown by subsequent activity.

Expanding on this class of non-linear problems, Hutchinson [10.6] and others developed a powerful approach and methodology for proceeding in general. Also, it should be observed that Rivlin and Thomas formulated fracture mechanics as explicitly applicable to highly non-linear elastomers.

Much contemporary work has centered around models of the non-linear conditions and the three-dimensional effects in the crack tip regions, especially at interfaces involved in the debonding of dissimilar materials. All these developments and contributions aggregate to a complete and widely used methodology for treating crack instability failure.

10.2 The Two Distinct Failure Theories

Fracture mechanics thus provides a highly useful method for approaching and solving many failure problems. Alternatively, many sources as well as this book show that failure criteria also provides a viable and comprehensive method for solving many failure problems. How should one approach the question of which formulation to use with a particular problem? One key to answering this question lies with the concept of homogeneity, which was extensively discussed in Section 2.2.

At some sufficiently large scale, most of the standard materials classes are taken to be homogeneous. Usually this scale is of the same order as the dominant scale of the applications of interest. At much smaller scales a new and vivid landscape of flaws, defects, and irregularities are to be seen. But at the macroscopic scale of application there exists a complete suite of homogeneous material properties—the intrinsic mechanical and thermal properties—that control behavior at this most common scale.

It is at the common scale of homogeneity of the material that most failure problems are most usefully posed. Of course, the macroscopic strength is profoundly affected by the subscale state of flaws, but not just by a single, idealized one of them. There is a whole distribution of them contributing to and causing failure at the macroscopic scale of the homogeneous material.

However, there is a second key needed to facilitate the comparison of the two failure theories. There is another independent scale of relevance and control that must also be considered. This scale emerges only in the particular problem of the intended application. This scale is that which implicitly governs the gradients of the applied stress state. If the gradients of the stress state (relative to the scale of the homogeneity of the material) are great, then fracture mechanics may be called for. If these stress gradients are shallow to moderate, then failure criteria are in order.

In the above context, failure criteria are defined and derived for the failure of homogeneous materials under homogeneous stress states. In effect failure, criteria represents the completion of the constitutive specification for the homogeneous material. Of course, the failure criteria and the other parts of the constitutive equations are then used in applications involving stress gradients so long as these gradients are not extreme. This is the simplest and most straightforward statement of the basis for the failure criteria methodology.

Is there an arbitrarily sharp division of scale between these two conditions of applicability? Not likely. But it is usually fairly obvious which case is present in a given problem. Stress conditions around sharp cracks and corners belong to the first group. Most other problems without extreme geometric curvature features fall into the second group. Thus it is both the scale of the homogeneity of the material and the scale (gradient) of the possible inhomogeneity of the stress state—extreme versus moderate—that helps decide which theory to use.

Possibly all this makes the decision process seem more complicated than it should be or actually is. Good judgment usually is enough to render the decision. Two examples will now be given—one where fracture mechanics is obviously the correct approach, and one where the use of failure criteria is clearly called for. Interest here is only with examples representing important, practical applications, not obscure problems of no real relevance or of misleading status.

10.3 Fracture Mechanics Example

The problem illustrating the application of fracture mechanics is that of the critical size of an edge crack in a load-carrying structural member. Typically, cracks form and grow near the edges and surfaces that are

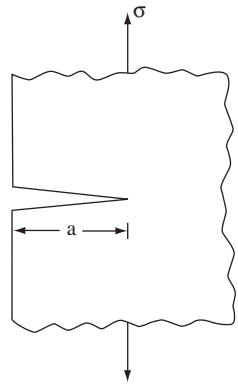


Fig. 10.1 *Edge crack.*

formed in processing and fabrication stages. The explicit problem is that of an edge crack, as shown in Fig. 10.1.

Often, cracks such as these grow under low level fatigue conditions. The question is, at what crack size a does the crack under constant load become unstable and cause failure? The governing fracture mechanics form is (10.6), and for this problem of the edge crack it is given by the stress intensity factor,

$$K_I = 1.12\sqrt{\pi}\sigma\sqrt{a} \tag{10.9}$$

To pose a specific problem, take the working stress level as

$$\sigma = 200\ MPa$$

For aluminum the critical stress intensity factor is about

$$K_{Ic} = 25\ MPa\sqrt{m}$$

Combining these last three equations gives the critical crack size a as

$$a = 3.97\ mm$$

This result is the critical crack size as found from linear elastic fracture mechanics. For significant plasticity behavior, as would be expected in

this problem, the corresponding critical crack size would be less than this value, thus

$$a < 4\ mm$$

So it is found that the critical crack size for this problem would certainly be less than 1 cm, and probably much less.

Although the explicit fracture mechanics formulas have been used here to estimate the critical crack size for a particular problem, that is not how it would normally be achieved in practice. Usually there are industry standards for the maximum allowable crack size in particular classes of problems, and these have been established from extensive databases. Nevertheless, this latter procedure still represents an explicit and critical use of fracture mechanics. In nearly all safety-related applications there are very comprehensive programs and protocols for detecting dangerous cracks and flaws and damage in order to prevent sudden fracture mechanics types failures.

10.4 Failure Criterion Example

The problem illustrating the necessary use of failure criteria is that for the single most basic problem in composite materials technology. This problem is that of the single very stiff and strong spherical inclusion in an elastic medium under far-field stress conditions—the dilute suspension case. This is the fundamental problem for the effective stiffness problem in elasticity as well as the effective viscosity for dilute fluid suspensions. The strength problem in the elasticity context is the complement of the effective stiffness problem.

To carry out this strength analysis, the three-dimensional elasticity solution is needed for the infinitely stiff spherical inclusion in the infinite elastic medium under far-field uniaxial tensile stress. Fig. 10.2 shows the problem of interest.

Finding the exact elasticity solution is not a trivial exercise. The stresses in the elastic medium are found by solving the governing coupled partial differential equations following a similar method given by Christensen [10.7] for the effective shear stiffness of the same constituents.

Perfectly continuous interface conditions are assumed. In general the stresses are a maximum at the interface with the spherical inclusion of

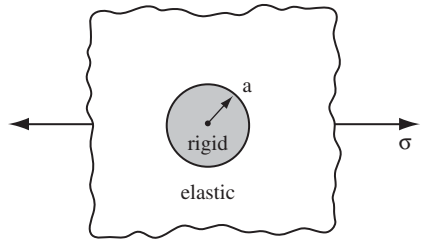


Fig. 10.2 *Rigid spherical inclusion in an infinite elastic medium.*

radius a. In spherical coordinates with the $\theta = 0$ axis being in the direction of the far-field uniaxial stress, σ, then at $r = a$ the stresses are found to be

$$\begin{aligned}
\sigma_r &= \frac{3\,(1-\nu)}{(4-5\nu)}\left[\frac{3}{(1+\nu)} - \frac{5}{2}\sin^2\theta\right]\sigma \\
\sigma_\theta = \sigma_\phi &= \frac{3\nu}{(4-5\nu)}\left[\frac{3}{(1+\nu)} - \frac{5}{2}\sin^2\theta\right]\sigma \\
\sigma_{r\theta} &= -\frac{15\,(1-\nu)\,\sigma}{2\,(4-5\nu)}\sin\theta\cos\theta
\end{aligned} \tag{10.10}$$

The other two stress components vanish identically because of symmetry.

Often in three-dimensional elasticity solutions the results take especially simple forms when Poisson's ratio has the value $\nu = 1/5$. In this case (10.10) becomes

$$\begin{aligned}
\sigma_r &= 2\sigma\cos^2\theta \\
\sigma_\theta = \sigma_\phi &= \frac{\sigma}{2}\cos^2\theta \\
\sigma_{r\theta} &= -2\sigma\sin\theta\cos\theta
\end{aligned} \tag{10.11}$$

The value $\nu = 1/5$ represents that for some glass and ceramic materials.

In composite materials, very stiff and strong inclusions are often placed within a polymeric matrix phase. The strength problem with such rigid inclusions therefore will be taken for a polymeric matrix material as a standard epoxy resin with typical properties of

$$\nu = \frac{2}{5} = 0.4$$

$$\frac{T}{C} = \frac{1}{2}$$

These properties and the stresses (10.10) at $r = a$ and at $\theta = 0$ will be used in the general, isotropic material failure criterion, eq. (4.17) repeated here as

$$\left(1 - \frac{T}{C}\right)(\hat{\sigma}_1 + \hat{\sigma}_2 + \hat{\sigma}_3) + \frac{1}{2}\left[(\hat{\sigma}_1 - \hat{\sigma}_2)^2 + (\hat{\sigma}_2 - \hat{\sigma}_3)^2 + (\hat{\sigma}_3 - \hat{\sigma}_1)^2\right] \leq \frac{T}{C} \tag{10.12}$$

The principal stresses in (10.12) are normalized by the uniaxial compressive strength C. This procedure then gives the strength result for the problem of Fig. 10.2 as

$$\sigma = 0.428\ T$$

Thus the local failure occurs, and in this brittle circumstance probably leads to overall failure, when the applied far-field uniaxial tensile stress is between 1/3 and 1/2 the value of the tensile strength of the polymeric matrix material. This is the effect of the stress concentration caused by the rigid inclusion. It can be obtained only quantitatively through the failure criterion, since more than one component of stress is operative.

In general, very stiff inclusions enhance the effective stiffness of the carrier matrix material. However, contrary to popular belief, in some cases and in this particular example the strength properties are degraded by the presence of the "reinforcing" inclusion(s). The problem becomes a trade-off between stiffness and strength.

10.5 Assessment

These two examples are typical of a great many realistic situations. There is a large array of very important problems that are covered by failure criteria, but there is an equally large and important collection covered by fracture mechanics. Neither approach can be said to be more important

than the other. They both are vitally important, and the two fields constitute complementary approaches in solving and codifying the critical failure problems involved in materials applications.

So there are two independent methodologies for dealing with failure, fracture mechanics, and failure criteria. The basic properties for both fields are needed in order to completely characterize the performance capability for any particular material in any particular application. It is interesting that there are parallel features shared by both of them. Fracture mechanics includes both brittle and ductile fracture behaviors, while failure criteria accommodates both ductile flow and brittle (and ductile) fracture for different materials types in different regions of stress space. To this extent there is an overlap between the two approaches. However, the differences are far greater than this superficial similarity.

Failure criteria, as developed here, is the rigorous theory of failure behavior for homogeneous materials under quasi-homogeneous (not extremely inhomogeneous) stress states. Fracture mechanics is an equally rigorous theory of failure behavior for the failure of structures (sometimes very simple or very small-scale structures) that always include a region of an extremely inhomogeneous stress state, the stress intensity zone, surrounding a crack or crack-like boundary condition. The term "structures" is used here rather than "materials" because boundary value problems are involved in the determination of the stress intensity factors. Both theories are often used beyond the narrow range of their derivations, and this becomes a matter of experience and judgment in applications.

Will there ever be a complete, general theory of failure that subsumes both of these approaches? It may be possible, but it is very unlikely. Attempting such a unified development would be an exceedingly difficult undertaking. There is great utility as well as considerable beauty in these two carefully constructed, carefully circumscribed, simpler theories: fracture mechanics and failure criteria.

Problem Areas for Study

1. Can failure criteria be formulated as a variational theory formalism? There has been some initial work in this direction, particularly for fracture mechanics. Can this approach be brought to a utilitarian level of application for failure criteria?

2. In fields with stress gradients, over what range can point-specific failure criteria be applied? Is it possible to use point-specific criteria over a region? If so, what would be the scale of the region and how would it relate to the stress gradient?
3. The path-independent J integral was developed by Rice for application to fracture mechanics. Is it conceivable that a path-independent integral could be developed for application to two-dimensional failure criteria?
4. Will there ever be a unified theory of both failure criteria and fracture mechanics? What are the pros and cons of even expending time and energy to pursue such a unified theory?

References

[10.1] Kanninen, M. F. and Popelar, C. H. (1985). *Advanced Fracture Mechanics*, Oxford University Press, New York.

[10.2] Broberg, K. B. (1999). *Cracks and Fracture*, Academic Press, New York.

[10.3] Suo, Z. (2010). *Fracture Mechanics*, http://imechanica.org/node/7448.

[10.4] Anderson, T. L. (2005). *Fracture Mechanics: Fundamentals and Applications*, 3rd ed., Taylor and Francis, Boca Raton.

[10.5] Rice, J. R. (1968). "A Path Independent Integral and the Approximate Analysis of Strain Concentration by Notches and Cracks," *J. Applied Mechanics*, **35**, 379–86.

[10.6] Hutchinson, J. W. (1983). "Fundamentals of the Phenomenological Theory on Nonlinear Fracture Mechanics," *J. Applied Mechanics*, **50**, 1042–51.

[10.7] Christensen, R. M. (2005). *Mechanics of Composite Materials*, Dover, New York.

11
Anisotropic Unidirectional Fiber Composites Failure

The anisotropic materials types to be considered here will be taken to have transversely isotropic symmetry. Perhaps the best-known example is that of aligned fiber composite materials, but there are many other examples. A further condition will be taken such that the degree of anisotropy is large. This is in line with the interests here in high-stiffness and high-strength fiber composite materials as typified by carbon fiber, polymeric matrix systems. Such materials will be referred to as carbon-polymer systems.

It is necessary to deduce the proper scale for the corresponding idealization of homogeneity for this class of materials failure problems. There are three obvious choices. The so-called micromechanics level takes the individual fibers and the separate matrix phase between them as the size scale for homogeneity. The next level up is the aligned fiber, lamina level, which then is much larger than the size of the individual filament or fiber. Finally, at yet a still much larger scale, the homogenization could be taken at the laminate level, involving the stacking of various lamina in various directions. It is the intermediate scale, the lamina level that is seen as having the proper balance between small-scale detail, but large enough scale to include all the possible failure mechanisms which could be operative. An example of the importance of the scale of the failure mode will be given later. Thus all idealizations to follow are taken at the aligned fiber, lamina scale of homogenization. This is the same scale as that at which the volume-averaged elastic properties for fiber composites are normally rationalized.

11.1 Transversely Isotropic Polynomial Invariants

The main purpose here is to develop the highly anisotropic failure criterion (for carbon-polymer systems) that is the companion piece to that of the isotropic case given in Chapter 4. To this end, take a polynomial

expansion of the stress tensor through terms of second degree for transversely isotropic symmetry. Such an expansion will involve the following seven terms composed of the four basic invariants for this symmetry and the three quadratic combinations of the two linear invariants,

$$\begin{aligned} &\sigma_{11,}\,(\sigma_{22}+\sigma_{33}),\ \sigma_{11}^2,\ (\sigma_{22}+\sigma_{33})^2, \\ &\sigma_{11}(\sigma_{22}+\sigma_{33}),\ (\sigma_{23}^2-\sigma_{22}\sigma_{33}),\ (\sigma_{12}^2+\sigma_{31}^2) \end{aligned} \tag{11.1}$$

Axis 1 is the axis of symmetry (fiber direction) and axes 2 and 3 form the plane of two-dimensional isotropy.

The condition of high anisotropy will be taken for both stiffness and strength, thus for the moduli

$$\frac{E_{11}}{E_{22}} \gg 1 \tag{11.2}$$

where E_{11} is the usual modulus in the fiber direction and E_{22} is the transverse modulus. For strengths, take T_{11} and C_{11} as the uniaxial tensile and compressive strengths in the fiber direction, and T_{22} and C_{22} as the corresponding strengths in the transverse direction. The highly anisotropic strengths are specified by

$$\frac{T_{11}}{T_{22}} \gg 1 \tag{11.3}$$

and

$$\frac{C_{11}}{C_{22}} \gg 1 \tag{11.4}$$

The lamina level strength properties that are conventionally measured by standard tests are the six following as

$$T_{11},\ C_{11},\ T_{22},\ C_{22},\ S_{12}\ \textit{and}\ S_{23} \tag{11.5}$$

where the Ts and Cs are already defined, S_{23} is the transverse shear strength, and $S_{12} = S_{31}$ is the longitudinal or axial shear strength. The conditions of high anisotropy also imply that S_{12} and S_{23} are of much smaller size than are T_{11} and C_{11}.

With this terminology and the highly anisotropic stiffness and strength conditions, the failure criteria can now be formulated. The failure will be found to naturally decompose into two separate modes.

First consider the conceptual limiting case of rigid fibers. Insofar as the matrix phase is concerned, the physical state would be that of plane strain, or alternatively that of out-of-plane shear deformation. For these stress states the effective macroscopic stresses controlling failure in the matrix phase would be those of the stresses in the 2–3 plane and the out-of-plane shear stresses.

For the cases of interest here where the fibers are not rigid but still are very stiff, (11.2), the stress forms resulting in matrix failure are taken to be the same as those in the above limiting case. The physical rationale for this is as follows. The anisotropic moduli ratio varies as $0 \leq E_{22}/E_{11} \leq 1$. The 0 limit is the plane strain case, and the 1 limit is that of isotropy. Just as a value of $E_{22}/E_{11} = 0.9$ would have the isotropic case as a close and reasonable representation, so too the conjugate value of $E_{22}/E_{11} = 0.1$ would have the plane-strain form as a close and reasonable representation. The value $E_{22}/E_{11} = 0.1$ or even considerably smaller values are well descriptive of typical carbon-polymer systems.

11.2 The Matrix-Controlled Failure Criterion

Taking the terms in (11.1) with the 2–3 plane components of stress and the out-of-plane shear stress components then gives the polynomial expansion as being comprised of the second, fourth, sixth, and seventh terms in (11.1). This gives the matrix-controlled failure criterion as

$$\alpha(\sigma_{22} + \sigma_{33}) + \beta(\sigma_{22} + \sigma_{33})^2 + \gamma(\sigma_{23}^2 - \sigma_{22}\sigma_{33}) + \delta(\sigma_{12}^2 + \sigma_{31}^2) \leq 1 \quad (11.6)$$

Evaluating parameters α, β, γ, and δ in (11.6) to give failure calibrated by the strength properties in (11.5) then gives the resulting failure criterion as

$$\begin{aligned} &\left(\frac{1}{T_{22}} - \frac{1}{C_{22}}\right)(\sigma_{22} + \sigma_{33}) + \frac{1}{T_{22}C_{22}}(\sigma_{22} + \sigma_{33})^2 \\ &\quad + \frac{1}{S_{23}^2}(\sigma_{23}^2 - \sigma_{22}\sigma_{33}) + \frac{1}{S_{12}^2}(\sigma_{12}^2 + \sigma_{31}^2) \leq 1 \end{aligned} \quad (11.7)$$

For (11.7) to always have real roots it is necessary that $S_{23} \geq \frac{1}{2}\sqrt{T_{22}C_{22}}$. As with isotropic materials it is commonly found that $T_{22} < C_{22}$ from tests. In Section 13.2 micromechanics will be used to express S_{23} in terms of T_{22} and C_{22}, it is a necessary and vital part of this failure theory.

The most common evidence of matrix-controlled failure is that of the transverse cracking in an aligned fiber lamina. When such failure occurs in a single, isolated lamina it represents complete, final failure of the piece. But if such local failure occurs in a single lamina that is located within a laminate, then it represents a state of damage. The subject of damage and failure in laminates will be taken up in Chapter 12.

11.3 The Fiber-Controlled Failure Criterion

The form (11.7) controls one of the two modes of failure. The other possible mode of failure is that of fiber-controlled failure. It could be tempting to assume that this means that fiber failure itself is the limiting failure mechanism. Such is not usually the case however. The problem is far more complex and subtle than that. Consider, for example, the compressive failure due to stress in the fiber direction, σ_{11}. The formation of kink bands is usually the failure mechanism. With the high-fiber stiffness, relation (11.2), the compressive failure mechanism is kink band formation with almost no deformation in the fiber direction, but the kink mechanism causes high shear stress in the matrix phase. The kink band formation occurs at the lamina scale, not the smaller scale nor the larger scale. The kink occurs suddenly as an instability which has the stress σ_{11} as proportional to the axial shear modulus. Even though the basic mechanism is that of an instability involving matrix deformation, it is still legitimate to designate this as fiber-controlled failure, since the high fiber stiffness plays the essential role, and the large failure stress is certainly related to the large fiber modulus. The situation with axial tensile failure is similarly complex, with fiber breaks and fiber misalignment playing crucial roles. The Rosen model includes some of these effects in a two-dimensional idealization.

The second mode of failure necessarily involves the other terms in (11.1) that are not part of the matrix-controlled failure, (11.7). These terms are then those of

$$\sigma_{11},\ \sigma_{11}^2,\ \sigma_{11}(\sigma_{22} + \sigma_{33}) \tag{11.8}$$

The large degree of anisotropy in strengths, (11.3) and (11.4), dictates that the fiber-controlled failure envelope have stress states with σ_{22} and σ_{33} being much smaller than σ_{11}. With this condition, then the last term in (11.8) is negligible (but not vanishing) compared with the first two terms. Writing the first two terms in failure criterion form and evaluating the two parameters in terms of the strength properties in (11.5) gives the resultant failure criterion as

$$\left(\frac{1}{T_{11}} - \frac{1}{C_{11}}\right)\sigma_{11} + \frac{1}{T_{11}C_{11}}\sigma_{11}^2 \leq 1 \tag{11.9a}$$

or simply

$$-C_{11} \leq \sigma_{11} \leq T_{11} \tag{11.9b}$$

In contrast to the matrix-controlled failure situation, it is commonly found that $C_{11} < T_{11}$.

The decomposed failure criteria (11.7) and (11.9) are completely calibrated by the six standard strength properties in (11.5). The fact that the fifth term of the expansion in (11.1) involving $\sigma_{11}(\sigma_{22} + \sigma_{33})$ does not enter either of these failure criteria is not an assumption, but rather the result of the rigorous derivation.

The background on the development of failure criteria (11.7) and (11.9) is as follows. The general outline of the method given here is similar to that developed by Christensen [11.1]. The end result, however, is different. In the reference given, special assumptions were made which reduced the failure criteria to four or five property forms. With interest in generality, no such assumptions are used here, leaving the properties count as six. The paired failure criteria (11.7) and (11.9) are the complete failure criteria to be derived here at the lamina level.

It should be noted that for applications the fiber direction stress σ_{11} must be taken in the fiber direction in the deformed configuration, not the reference configuration. To do otherwise could cause σ_{11} to induce a very large longitudinal shear stress that would certainly cause matrix-controlled failure.

The physical significance of the present anisotropic failure criteria, (11.7) and (11.9), is that they are the direct counterpart of the isotropic failure criteria given in Section 4.4. As seen from Sections 4.4 and 4.5, failure in isotropic materials is much more highly developed than is the

anisotropic case. Nevertheless, as shown here, significant progress has been made in the more difficult anisotropic case.

The fiber-controlled failure form (11.9) has an interesting history. This form is commonly called the maximum stress criterion. It has always been considered to be a highly useful but totally empirical form for fiber composites. In the present derivation it is not empirical at all; it is a rational and rigorous result of the method, which is based upon the conditions of a high degree of anisotropy, with no subsidiary assumptions. In this connection it can also be observed that the matrix-controlled failure criterion, (11.7), certainly is not a maximum stress form. The two criteria (11.7) and (11.9) seem to be very different when compared through the commonly used form (11.9b), but not so greatly different when compared through the equivalent but more formal (11.9a). The two coordinated failure criteria, (11.7) and (11.9), are the end result of this physical derivation.

The failure criteria (11.7) and (11.9) are thus fundamentally based upon the high degree of anisotropy conditions (11.2)–(11.4), which in turn are motivated by the properties of carbon-polymer systems. Although some other types of fiber composites may not satisfy the high anisotropy conditions, they probably would still favor the separation of failure modes as in (11.7) and (11.9). Most systems at high fiber concentration have failure modes strongly influenced by the fiber-to-matrix morphology. As an example, transverse cracking as a matrix-controlled failure mode is common to virtually all fiber composites.

Two much earlier and well-known failure criteria for fiber composites are those of the Tsai–Wu form and the Hashin form. These two criteria will be stated here for comparison with (11.7) and (11.9).

11.4 Hashin Failure Criterion

The Hashin criterion [11.2] starts with the same seven terms as in (11.1), and then decomposes into separate fiber and matrix failure modes. It also distinguishes essentially tensile states from compressive states, with separate criteria taken for each. This then involves many more terms than just those in (11.1) being used singly, since some are used twice. Finally, several assumptions are used to bring the parameter count down to a manageable level of the six properties in (11.5). The end result is the Hashin criterion given by:

Tensile Matrix Mode, $(\sigma_{22} + \sigma_{33}) > 0$

$$\frac{1}{T_{22}^2}(\sigma_{22} + \sigma_{33})^2 + \frac{1}{S_{23}^2}\left(\sigma_{23}^2 - \sigma_{22}\sigma_{23}\right) + \frac{1}{S_{12}^2}\left(\sigma_{12}^2 + \sigma_{31}^2\right) \leq 1$$

Compressive Matrix Mode, $(\sigma_{22} + \sigma_{33}) < 0$

$$\frac{1}{C_{22}}\left[\left(\frac{C_{22}}{2S_{23}}\right)^2 - 1\right](\sigma_{22} + \sigma_{33}) + \frac{1}{4S_{23}^2}(\sigma_{22} + \sigma_{33})^2 + \frac{1}{S_{23}^2}(\sigma_{23}^2 - \sigma_{22}\sigma_{33}) + \frac{1}{S_{12}^2}(\sigma_{12}^2 + \sigma_{31}^2) \leq 1$$

Tensile Fiber Mode, $\sigma_{11} > 0$

$$\left(\frac{\sigma_{11}}{T_{11}}\right)^2 + \frac{1}{S_{12}^2}(\sigma_{12}^2 + \sigma_{31}^2) \leq 1$$

Compressive Fiber Mode, $\sigma_{11} < 0$

$$\left(\frac{\sigma_{11}}{C_{11}}\right)^2 \leq 1 \tag{11.10}$$

The Hashin criterion is thus composed of four separate modes of failure.

11.5 Tsai–Wu Failure Criterion

The Tsai–Wu criterion [11.3] combines all terms in (11.1) directly into a single mode of failure, and gives a seven-property form involving the six properties in (11.5), plus one other, F_{12} below. The Tsai–Wu criterion is:

$$\left(\frac{1}{T_{11}} - \frac{1}{C_{11}}\right)\sigma_{11} + \left(\frac{1}{T_{22}} - \frac{1}{C_{22}}\right)(\sigma_{22} + \sigma_{33}) + \frac{\sigma_{11}^2}{T_{11}C_{11}} + \frac{1}{T_{22}C_{22}}(\sigma_{22} + \sigma_{33})^2 + F_{12}\sigma_{11}(\sigma_{22} + \sigma_{33}) + \frac{1}{S_{23}^2}\left(\sigma_{23}^2 - \sigma_{22}\sigma_{33}\right) + \frac{1}{S_{12}^2}\left(\sigma_{12}^2 + \sigma_{31}^2\right) \leq 1 \tag{11.11}$$

The F_{12} parameter in (11.11) is usually called an interaction parameter. It must be specified by some auxiliary means. Often F_{12} is scaled differently by writing $2F_{12}$ rather than F_{12} in (11.11). Some published forms for the Tsai–Wu criterion also contain a second interaction parameter, F_{23}, in addition to the properties in (11.5), but that form is redundant, and the form given here without F_{23} must be used to be consistent.

11.6 Comparisons

It is apparent that the present failure criteria (11.7) and (11.9) are much different from the Hashin and the Tsai–Wu forms. Later examples will reveal strong differences between all three failure criteria for carbon-polymer systems. The Tsai–Wu form does not decompose into separate fiber- and matrix-controlled modes; whereas the other two criteria do so decompose, the present formalism has the two obvious modes of failure while the Hashin form has four modes. This question of the possible decomposition or not of failure modes has always been of central importance to the field, and one that has generated strong positions and debates. This will be discussed further below.

The main difference between the present forms and the Hashin forms is that no assumptions are involved with the present forms, whereas five particular assumptions were necessarily involved in the Hashin derivation. In addition, the Hashin method further decomposes the fiber- and matrix-controlled modes into submodes of either tensile or compressive nature. The isotropic material results of the previous section shows that this "subdecomposition" is unnecessary and inappropriate.

Following the method given here, under the condition of high anisotropy, the failure characterization must decompose into the two separate failure criteria. On this basis the Tsai–Wu form can only apply to moderately anisotropic systems; that is, only for moderate departures from a state of isotropy. It cannot recover the limiting plane strain condition. The other two criteria do not apply under such moderately anisotropic conditions, since they do not admit the limiting case of isotropy. They apply in the high-anisotropy case typified by carbon-polymer systems. Thus the decomposed failure mode forms apply near one end of the anisotropy scale and recover the plane strain condition, while the undecomposed form applies near the other end of the same scale and recovers the isotropy condition.

Considering the many composites failure criteria that have been proposed over time, the three general criteria discussed here stand out as having substantial derivations and developments. Of these three criteria, the present failure criterion, (11.7) and (11.9), has the most physically realistic basis for application to high-performance fiber composite materials. An example of other criteria is that of Puck and colleagues, given by Puck and Schurmann [11.4]. Hinton, Kaddour, and Sodden [11.5] conducted an evaluation exercise for fiber composite failure criteria that provides a broad and helpful view of related matters.

We now return to the use and interpretation of the failure criteria (11.7) and (11.9). In order to present examples it is necessary to assign elastic and failure property values. Typical reported properties for carbon-epoxy systems are given by

$$\begin{aligned} E_{11} &= 150\,\text{GPa} \\ E_{22} &= \quad 9\,\text{GPa} \\ \mu_{12} &= \quad 6\,\text{GPa} \\ \mu_{23} &= \quad 3\,\text{GPa} \\ \upsilon_{12} &= \quad 1/3 \\ \upsilon_{23} &= \quad 1/2 \end{aligned}$$

and

$$\begin{aligned} T_{11} &= 2000\,\text{MPa} \\ C_{11} &= 1500\,\text{MPa} \\ T_{22} &= \quad 40\,\text{MPa} \\ C_{22} &= \quad 150\,\text{MPa} \\ S_{12} &= \quad 80\,\text{MPa} \\ S_{23} &= \quad 50\,\text{MPa} \end{aligned}$$

where the shear moduli and Poisson's ratios are included. It is seen that these typical properties conform to the high degree of anisotropy in both stiffness and strength. See Section 13.2 for restrictions on S_{23}.

The matrix-controlled failure mode, (11.7), ordinarily involves curved failure envelopes such as the ellipses and parabolas of the isotropic case in the previous section. The most interesting cases are for stress states involving the fiber direction stress, σ_{11}, along with some of the other stress components, thus bringing in both failure modes, (11.7) and (11.9). Shown

in Figs. 11.1 and 11.2 are the two stress states: σ_{11} versus σ_{22}, and σ_{11} versus $\sigma_{22} = \sigma_{33}$—both cases with the other stresses as zero.

The three-dimensional case of Fig. 11.2 gives a transverse tensile value of about half that of the two-dimensional case in Fig. 11.1. Fig. 11.2 gives a transverse compressive value slightly larger in magnitude that that of Fig. 11.1. The Hashin form (11.10) gives the same result for the case shown in Fig. 11.1, but does not give a lower limit for $\sigma_{22} = \sigma_{33}$ in the case of Fig. 11.2. The Tsai–Wu form (11.11) in both cases gives slender elliptical envelopes that at the maximum extend considerably beyond the fiber-direction uniaxial strengths. Section 13.2 has other examples.

As a more involved example than those of Figs. 11.1 and 11.2, take the realistic stress state with $\sigma_{11} = 1500$ MPa ($^3/_4$ of the uniaxial tensile strength) and then establish the failure envelope for σ_{22} versus σ_{33}. The present criteria (11.7) and (11.9) give an elliptical envelope. The Tsai–Wu criterion also gives an elliptical envelope, but is considerably smaller

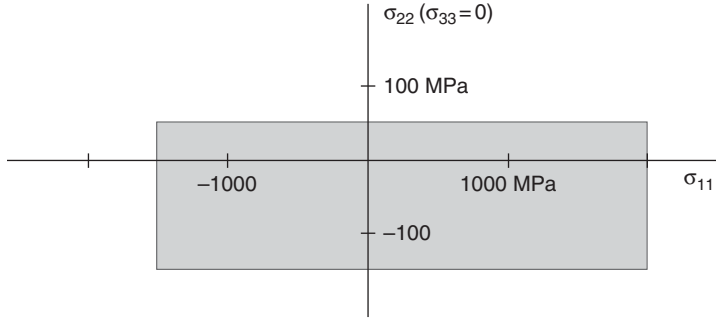


Fig. 11.1 *Two-dimensional stress state, (11.7) and (11.9).*

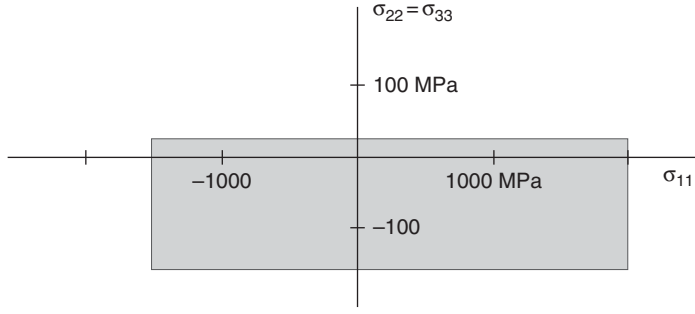


Fig. 11.2 *Three-dimensional stress state, (11.7) and (11.9).*

and shifted from than that from (11.7). The Hashin criterion gives an open-ended parabolic form. To be more specific in this example would require selecting a value for the F_{12} parameter in the Tsai–Wu form (11.11). Interested parties should work out cases such as these to show the differences between these three criteria, or any others. They all are fundamentally different.

The corners shown in Figs. 11.1 and 11.2 are the direct result of the decomposition into separate fiber-controlled and matrix-controlled modes of failure. This situation is completely parallel to that of the intersecting failure modes of isotropic materials in Chapters 4 and 5. In physical reality, the corners would be expected to be rounded due to the small misalignment of testing specimens, and due to existing states of damage and specific inhomogeneities, as well as a host of other non-ideal effects. To some extent this characteristic may also relate to the testing of lamina in isolation, as opposed to the stabilized behavior of an *in situ* lamina within a laminate. From this point of view it is certainly best to use the forms directly from the basic failure criteria, (11.7) and (11.9), rather than to combine them with some artificial smoothing technique. Another problem with the regions around corners in failure surfaces is the extreme difficulty in generating reliable multi-axial test data. It is very much more difficult than that for the one-dimensional strength tests of the basic properties in (11.5).

Fiber composites are rarely used in unidirectional form. Most commonly, lamina (or tows) are taken as the building blocks in laminates (or woven forms) composed of layers at various orientations. It is with laminates where the questions of corners in the failure surfaces are best considered and treated. Such topics will be taken up in the next section using the lamina level failure criteria (11.7) and (11.9) as the foundation and starting point.

Problem Areas for Study

1. For isotropic materials a comprehensive treatment of ductile-versus-brittle failure behavior has been developed. Are similarly general ductile-versus-brittle developments possible for fiber composite materials? If general developments are not forthcoming, at least could the

transverse tensile and fiber-direction tensile strengths admit a ductile-versus-brittle criterion? Their D/B behaviors presently are rationalized only on an intuitive basis!

2. Are unidirectional, aligned fiber composites the ideal or optimal way to combine two material phases to enhance both stiffness and strength?
3. Testing the validity of failure theories for fiber composites is under investigation in the ongoing World-Wide Failure Exercise (WWFE) conducted by M. J. Hinton and A. S. Kaddour. This large-scale evaluation program focuses upon the most difficult problems related to the load-carrying capabilities of fiber composites. Any current activities of this type should be compared with and correlated with past, present, and future investigations in the WWFE. Are there any means and mechanisms for assimilating all the fragmentary evidence already in existence, some real and some little more than hearsay?
4. Some aligned fiber composite material failure theories are explicitly expressed separately in terms of tensile failure modes and compressive failure modes. The present general failure theory for aligned fiber composites does not require degeneration and separate specification in tensile and compressive modes. Its quadratic form automatically yields the tensile and compressive modes in the same way as for isotropy. Are there any contradictions between the two different approaches?
5. Relative to any particular symmetry class, is there a relationship between the number of independent elastic properties and the number of independent failure properties? For isotropy there are two of each according to the methods given here. How do fiber composites compare in this regard?

References

[11.1] Christensen, R. M. (1997). "Stress Based Yield/Failure Criteria for Fiber Composites," *Int. J. Solids Structures*, **34**, 529–43; see also *J. Engr. Mats and Technology* (1998). **120**, 110–3.

[11.2] Hashin, Z. (1980). "Failure Criteria for Unidirectional Fiber Composites," *J. Appl. Mech.* **47**, 329–34.

[11.3] Tsai, S. W. and Wu, E. M. (1971). "A General Theory of Strength for Anisotropic Materials," *J. Comp. Mater.* **5**, 58–80.

[11.4] Puck, A. and Schurmann, H. (2002). "Failure Analysis of FRP Laminates by means of Physically Based Phenomenological Models," *Comps. Sci. and Technology*, **62**, 1633–62.

[11.5] Hinton, M. J., Kaddour, A. S., and Soden, P. D. (2002). "A Comparison of the Predictive Capabilities of Current Failure Theories for Composite Laminates, Judged against Experimental Evidence," *Comps. Sci. and Technology*, **62**, 1725–97.

12
Anisotropic Fiber Composite Laminates Failure

Fiber composites are typically used in flat, laminated forms or variations thereof, often also involving woven or braided constructions. The treatment given here will be for the flat laminated form where the aligned fibers in the various lamina take different directions within the laminate. Failure at the laminate level is complex and not an obvious matter in terms of deducing the theoretical basis for the related failure criteria. Two very different but mainline methods will be presented here.

12.1 Introduction

The best place to start is with an understanding of the defect states that can and do exist in commercial fiber composite products. These defects have a profound influence upon the strength performance of the resulting materials. There are broken fibers, fiber slacks, fiber misalignments, debonds, resin-rich pockets, cracks, porosity, and on and on. Effective moduli do not depend much on the defects, but failure surely does.

With all the defects, one could question how any composite material could perform at a satisfactory level. Despite the many defects, the special properties of the fibers and the ameliorating role of the matrix combine to produce an extraordinary balance of properties that remains at a very high level for the higher-quality products. To be sure, these composite properties would not be possible without a superior capability of the fiber phase or the enabling properties of the matrix phase, as both are vital and essential.

Next on the scale of importance is the matter of scale itself. One of the major fulcrums of dissension is the uncertainty as to the scale at which to characterize failure. All failure modes have a scale of action. The central problem is to recognize the scale for a particular mode of failure

and to then characterize it. Damage occurs at all scales from molecular on up.

Perhaps an extreme example will best illustrate the point. Molecular dynamics cannot predict the Euler buckling of a major structural column. The mismatch in scales is ridiculously great. While it is generally taken that smaller scales reveal more fundamental effects than do larger scales, there are important and critical exceptions to this oversimplification. Many or most materials composing load-bearing structures fall under the exceptional category. These engineering materials are strongly distinguished from other major classes such as electronic materials and bio-materials which have far different primary functions from that of mainly supporting load without failure but with minimal deformation.

Composites laminates have failure modes that inherently relate to the scale of the laminate. For example, delamination is meaningful only at the laminate scale. The different but very broad scales of conceivable relevance here are

Atomic—angstroms 10^{-9} m
Fiber size—microns 10^{-6} m
Lamina thickness—mils 10^{-4} m
Laminate thickness—millimeters to inches 10^{-1} m

Concerning scale dependence, kink bands, for example, cannot be predicted from fiber scale failure. More generally, contrary to many claims, macroscopic failure cannot be predicted by molecular dynamics. There is a six or seven orders-of-magnitude gap that is an imposing barrier. Realistic macroscopic failure predictions must originate from a careful assessment of the proper scale for the failure mode. Many, but not necessarily all, macroscopic failure modes originate at the macroscopic scale. Great care is needed in making this determination. Taking macroscopic failure to be at the lamina and laminate scales, the origin of the failure modes then depends upon the particulars of the layup of the lamina within the laminate.

Although interesting as concepts and certainly useful in developing new materials, truly using nanomechanics and/or micromechanics effects to predict macroscopic failure would require an immense commitment, and by no means is it an established approach. Some microscale or nanoscale approaches are that in name only, since they "back out" constituent

properties to produce desired macroscopic scale results. In fact, even just trying to predict laminate failure behavior from lamina level behavior may have severe limitations, as will be shown later. Due to the complications of damage and of the scale dependence of the failure modes, failure characterization based upon nanoscale and microscale idealizations (either fully or partially so) will not be considered in the two methods to be presented in this chapter.

Even though the work in this chapter and the preceding chapters is focused upon macroscopic failure with the corresponding macroscale level of analyses, this does not mean that microscale and nanoscale results are not useful. Of course, nanoscale and microscale investigations are vitally important. The next two chapters, 13 and 14, will be used to demonstrate the tremendous usefulness and insights to be gained from microscale and nanoscale investigations in some cases.

There is a very large literature on the subject of damage and failure in fiber composite laminates. All conceivable approaches are being tried, and a few typical approaches will be mentioned here. Hinton, Kaddour, and Soden [12.1] compiled a very useful collection and assessment of many different approaches; Puck and Schurmann [12.2] present an approach originally based upon Coulomb–Mohr-type behavior; Mayes and Hansen [12.3] descriptively designate their method as MCT (multicontinuum theory); Daniel [12.4] treats many aspects of failure; Robbins and Reddy [12.5] present an approach using internal variables; Tsai and colleagues [12.6] present and use MMF (micromechanics of failure); and a finite element approach to damage and failure, EFM (element failure method), is given by Tay *et al.* [12.7]. Not only are there many research sources and references, but almost all of them represent distinctly different and individualized approaches. Considering the difficulty of the topic, it is not surprising that there is nothing that is even within proximity of being a unified, verified, and reasonably recognized methodology.

The present interests and methods will be restricted to carbon fiber–polymeric matrix systems, which are generally the highest-performance systems. As such, these approaches may be somewhat less applicable to glass–polymer systems, although the possibility remains open.

The technical starting point is to recall the lamina-level treatment given in Chapter 11. Typical properties for aligned carbon fiber–polymer matrix lamina are given by

$$
\begin{aligned}
\mathrm{E}_{11} &= 150\,\mathrm{GPa} \\
\mathrm{E}_{22} &= \quad 9\,\mathrm{GPa} \\
\mu_{12} &= \quad 6\,\mathrm{GPa} \\
\mu_{23} &= \quad 3\,\mathrm{GPa} \\
\nu_{12} &= 1/3 \\
\nu_{23} &= 1/2
\end{aligned}
$$

and

$$
\begin{aligned}
\mathrm{T}_{11} &= 2000\,\mathrm{MPa} \\
\mathrm{C}_{11} &= 1500\,\mathrm{MPa} \\
\mathrm{T}_{22} &= \quad 40\,\mathrm{MPa} \\
\mathrm{C}_{22} &= \quad 150\,\mathrm{MPa} \\
\mathrm{S}_{12} &= \quad 80\,\mathrm{MPa} \\
\mathrm{S}_{23} &= \quad 50\,\mathrm{MPa}
\end{aligned}
\tag{12.1}
$$

The two independent lamina-level failure criteria derived in Sections 11.1–11.3 are given by:

Fiber-Controlled Failure

$$
\left(\frac{1}{T_{11}} - \frac{1}{C_{11}}\right)\sigma_{11} + \frac{1}{T_{11}C_{11}}\sigma_{11}^2 \leq 1 \tag{12.2a}
$$

or simply

$$
-C_{11} \leq \sigma_{11} \leq T_{11} \tag{12.2b}
$$

Matrix-Controlled Failure

$$
\begin{aligned}
&\left(\frac{1}{T_{22}} - \frac{1}{C_{22}}\right)\left(\sigma_{22} + \sigma_{33}\right) + \frac{1}{T_{22}C_{22}}\left(\sigma_{22} + \sigma_{33}\right)^2 \\
&\quad + \frac{1}{S_{23}^2}\left(\sigma_{23}^2 - \sigma_{22}\sigma_{33}\right) + \frac{1}{S_{12}^2}\left(\sigma_{12}^2 + \sigma_{31}^2\right) \leq 1
\end{aligned}
\tag{12.3}
$$

There is something especially attractive about a lamina-level failure criterion that decomposes into separate fiber-controlled and matrix-controlled failure modes. They do not constitute the obvious approach of explicitly treating fiber failure as distinct from matrix failure and conversely treating matrix failure as distinct from fiber failure. They do allow

interactive fiber and matrix effects but with all the defects included in an implicit way, and still they decompose into the two separated modes of failure. In Chapter 11 two other lamina-level failure criteria are presented: namely, the Tsai–Wu and the Hashin forms. The first does not allow the decomposition of failure modes, while the second involves four different decomposed subforms.

Concerning the uniaxial lamina-failure criteria given above, it is seen that the stress in the fiber direction is uncoupled from the other stress components. Should all these stress components be coupled? One could assume a coupled form and then evaluate the coupling parameters from failure data. While these would then fit a particular set of data, they could hardly be considered to be general properties of failure. It is here preferred to use the theoretically derived and idealized form specified by only the explicit failure properties in the above criterion, and by nothing else. This, then, is considered to give the essential elements of failure at the lamina scale. This is what will be used in approaching the method of progressive damage at the laminate scale. Actually, this method bridges between the lamina and laminate scales.

12.2 Progressive Damage in Laminates

The idea behind progressive damage is quite simple. Both matrix-controlled and fiber-controlled types of failure can separately and sequentially occur during the loading of the various lamina within a laminate. At some point so much damage has accumulated in the form of these local failures that the laminate can no longer sustain load. This, then, comprises the ultimate load—failure in the broad and total sense.

Progressive damage in laminates, therefore, is not so much an original and stand-alone discipline as it is a careful accounting of the sequence of local failures (damage) leading to the complete failure of the laminate. The lamina-level failure criteria for the fiber-controlled and the matrix-controlled modes still must be selected from a variety of existing forms. Thus the term "progressive damage" essentially means predicting laminate failure from lamina-level damage and failure.

The standard, direct way of proceeding places high priority on modeling the failure at the lamina level and then using that to build up to laminate failure. The presumption then is that high accuracy at the lamina level would translate into highly reliable results at the laminate level.

This is the conventional thinking. While this line is quite reasonable on its face, the problem is in what constitutes high accuracy at the lamina level, and how that translates to the laminate level.

In the immediately following developments, the in-plane and the out-of-plane failure modes for the laminate will be taken to be uncoupled due to the extreme anisotropy of effects between them. First, the in-planes cases will be examined in considerable detail, and later the out-of-plane (delamination) cases will be taken up.

We begin with failure of the quasi-isotropic layup case. The common form for this involves equal numbers of lamina in the 0, +45, –45, and 90-degree fiber directions. The failure of the quasi-isotropic layup is probably the most important single case that can be studied. This limiting case is one of the two anchors of all possible layups, the other limiting case is the unidirectional form itself, and all other cases lie between these two extremes. Thus the quasi-isotropic case is one of the most severe tests of using a lamina-level failure criterion to predict damage and failure of a laminate.

The first damage that usually occurs is that of the matrix failure in some lamina. But this is not of as much interest here as the complete failure involving the fiber load-carrying capability. Taking the properties and lamina level failure criteria given in (12.1), the exact sequence of failures can be determined for the various lamina within the quasi-isotropic laminate. These are given by the following sequences for some particularly important stress states. The unidirectional stress states are aligned with the 0-degree fiber direction. Lamina-level failure criteria (12.2) and (12.3) along with properties (12.1) are used in the progressive damage analysis.

Uniaxial tension
- (i) 90-degree matrix damage, then
- (ii) ±45-degree matrix damage, then
- (iii) 0-degree fiber failure

Uniaxial compression
- (i) ±45-degree matrix damage, simultaneous with
- (i) 0-degree fiber failure

Eqi-biaxial tension
- (i) All matrix damage, then
- (ii) All fiber failure

Eqi-biaxial compression
  (i) All fiber failure, no matrix damage
Shear (0-degree tension, 90-degree compression)
  (i) ±45-degree matrix damage, then
  (ii) 90-degree matrix damage, then
  (iii) 90-degree fiber failure

In these examples the corresponding moduli are set to zero after the failure criteria prescribes the matrix-controlled or fiber-controlled failure. The fiber failures shown lead directly to overloading the other lamina, and total failure ensues. The details of this accounting approach are lengthy, but completely routine.

When the failure mode is matrix controlled, the resultant "jump" in strain at constant stress is small compared with the jump in strain that occurs when the fiber mode fails in a particular lamina. This motivates a simplification that is here designated as "fiber-dominated" progressive damage. Again, this is appropriate mainly to carbon-polymer systems, and it preserves the general shape of the failure envelopes.

Henceforth in the progressive damage examples the fiber-dominated state will be used whereby the matrix-controlled properties are taken to be negligible compared with the fiber-controlled properties. Then only a sequence of fiber-controlled failures in the various lamina need be considered.

For the present properties example, and for the quasi-isotropic laminate, the fiber-dominated biaxial stress failure envelope is given in Fig. 12.1 for the case of $T = C$. In this case subscripts are not needed on the in-plane uniaxial strengths. Tensile and compressive strengths of 2000 MPa are used for the fiber-direction lamina strengths.

In Fig. 12.1 and from this point on, the 1, 2, 3 coordinate system refers to the laminate with the 3-axis in the thickness direction. It is seen that the failure envelope takes a diamond-shaped form. There is reflection symmetry about axes that are rotated 45 degrees to those in Fig. 12.1.

The failure envelope due to progressive damage in the quasi-isotropic laminate with fiber-direction strengths from (12.1) is shown in Fig. 12.2. The $T \neq C$ case has a lower degree of symmetry than that for $T = C$. The diamond shape persists and is characteristic of the progressive damage approach.

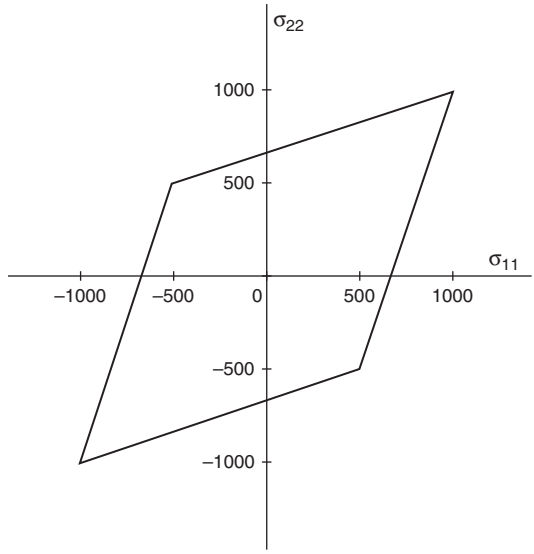


Fig. 12.1 *Quasi-isotropic failure envelope, progressive damage, $T = C$.*

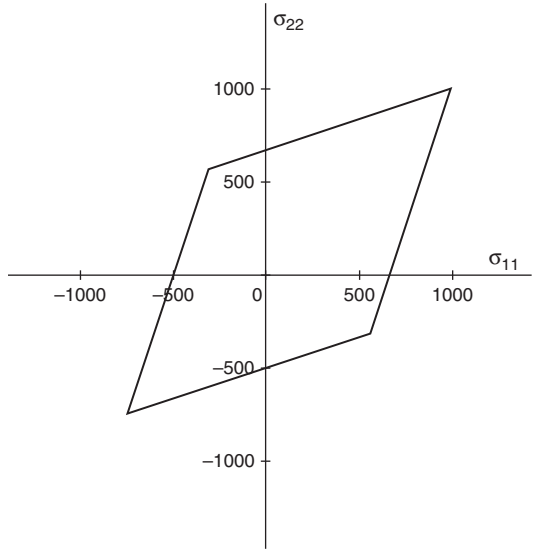


Fig. 12.2 *Quasi-isotropic failure envelope, progressive damage, $T \neq C$.*

Now an orthotropic example will be given. Take a layup that has a strong emphasis on the reinforcement in a particular direction, with the volume fractions

$$\begin{aligned} 0^\circ &\text{ lamina, } c = 1/2 \\ +45^\circ &\text{ lamina, } c = 1/6 \\ -45^\circ &\text{ lamina, } c = 1/6 \\ 90^\circ &\text{ lamina, } c = 1/6 \end{aligned} \tag{12.4}$$

Using the fiber-dominated properties from (12.1) gives the failure envelope as shown in Fig. 12.3, where it can be seen that the degree of anisotropy in the strengths is larger than 2 to 1.

These diamond-shaped failure envelopes are like those of the maximum normal strain criterion for three-dimensionally isotropic materials. This similarity is not just a coincidence. At the lamina level the failure criterion is that of stress in the fiber direction. For a high degree of anisotropy at the lamina level the criterion of stress in the fiber direction is well approximated by strain in the fiber direction. This interpretation is useful, because strain in the fiber direction at the lamina level is the same as strain in a fiber direction at the laminate level, for the in-plane conditions.

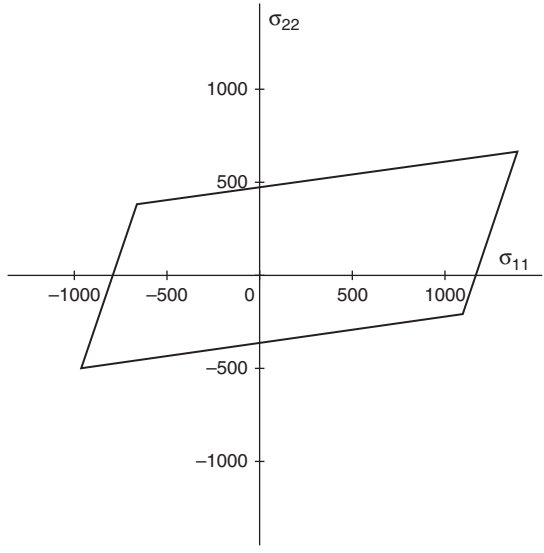


Fig. 12.3 *Orthotropic failure envelope, progressive damage.*

Thus in these examples the laminate failure criterion is that of maximum strain in a (any) fiber direction.

The progressive damage behavior shown in Figs. 12.1–12.3 is typical for most layup forms. These results are somewhat surprising, especially as regards the acute angles of the failure envelopes in the first and third quadrants and the associated very large allowable stresses. The obvious question concerns whether these acute angles reflect expected physical behavior or whether they are artifacts of the method of progressive damage.

12.3 Testing Results

Hinton, Kaddour, and Soden [12.1] have conducted a lengthy and detailed evaluation of many different fiber composite failure theories. Most of the theories that were included in their study involved variations of progressive damage. In particular, for a quasi-isotropic laminate they found rather good agreement between data that they cited (by Swanson) and progressive damage in the first, second, and fourth quadrants of biaxial stress states. It should be noted that the data in the first quadrant did not extend all the way to the apex of the predicted failure locus. They did not find good agreement in the third quadrant, but that was attributed to an experimental deficiency, the buckling of the specimens in compression–compression stress states.

There is a different source of data that gives a different conclusion from that mentioned above. Specifically, data generated and reported by Welsh, Mayes, and Biskner [12.8] does not support the first and third quadrant predictions of progressive damage for carbon-epoxy, quasi-isotropic laminates in biaxial stress states. In fact, they reported that their data are fairly well described by a Mises-like form with no corners at all in the failure envelope.

The Welsh, Mayes, Biskner data are shown in Fig. 12.4, using symmetry about a 45-degree line through the origin.

Thus there are conflicting datasets for the failure behavior of quasi-isotropic laminates. One dataset supports the existence of the acute angle corners in the biaxial failure envelope, while another does not support it.

From a theoretical point of view, the form with corners is prescribed by the maximum strain type of failure criterion (in a fiber direction). The maximum strain criterion has no credibility for isotropic materials such

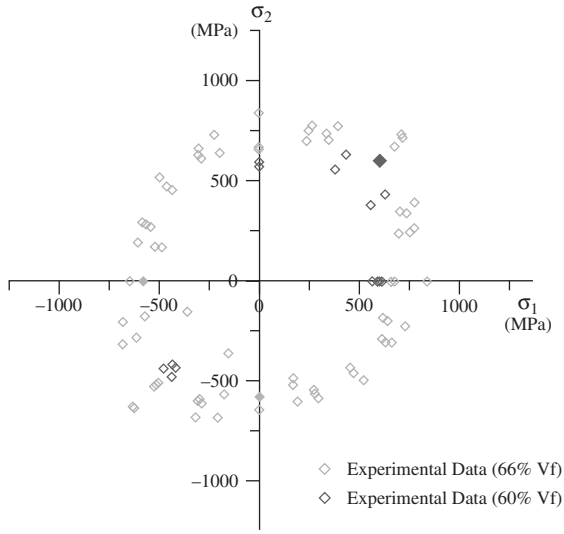


Fig. 12.4 *Quasi-isotropic laminate failure data, IM7/977-2 (courtesy of J. S. Welsh).*

as metals and polymers, but despite this, does it and should it apply for fiber composite laminates? The entire situation is somewhat ambiguous because of its complications.

Perhaps the strain criterion in the fiber direction is too simplistic for laminates. It is a one-dimensional criterion. Perhaps laminates, as distinct from a lamina within the laminate, requires a failure criterion that allows more than one-dimensional stress and strain representation. Perhaps laminates allow more critical (lower-level) failure modes than those permitted by one-dimensional strain. The genesis for the one-dimensional strain criterion is probably the overly simplified view that composites are composed of perfectly straight fibers that are perfectly collimated and spaced in an ideal matrix phase.

It is seen that this progressive-damage approach is built upon a set of conjectures as to what the failure mode should be. However, a qualification should be added to this. It is best viewed as what the failure mode "could be," not "should be." It may be best to view this standard progressive-damage approach as an upper bound. It is a plausible failure mode, but it could be superseded by a different failure mode that operates at a different scale and at a lower critical value of stress in the laminate.

Another complication is that of *in situ* properties and damage. In quasi-isotropic laminates under uniaxial tensile stress the lamina under the most load in the fiber direction has no significant transverse stresses causing matrix cracking. But in eqi-biaxial tension all lamina have very extensive matrix cracking and damage. The transverse cracking also can further degenerate into delamination. In eqi-biaxial tension the damage is invasive, and compromises the load-carrying capacity in the fiber direction within any lamina. The diamond-shaped failure envelope from progressive damage is too idealized to accurately reflect such degrading effects. A more conservative and more physically realistic failure envelope may be called for, and is examined next.

12.4 Polynomial Invariants for Laminates

If there were a need to characterize failure for a given laminate configuration and the method of progressive damage were not available, how would one approach the problem? Almost certainly one would use the same general method as was used in Chapter 4 for isotropic materials. Specifically, one would take a polynomial expansion in the stress invariants for the symmetry of interest, and then truncate the expansion at second-degree terms. As then expressed, the form would involve specific constants or parameters to be interpreted as the failure properties. By obvious description the method should be designated as that of "polynomial invariants." Exactly this method will be developed and followed here for the failure of laminates.

The reason why the polynomial-invariants method may be more appropriate than progressive damage is that some failures are at the laminate level, not the lamina level. Thus some failure modes involve cooperative motions between the lamina within the laminate. Any failure mode has a scale of action at which it is relevant, and the laminate failure modes should be considered.

The first developments and examples will be for quasi-isotropic laminates. The failure modes will be concerned with the in-plane ones, and then later the out-of-plane ones (delamination) will be developed. The term quasi-isotropic refers to in-plane elastic moduli (or compliance) properties. It is an entirely separate question to consider whether the failure properties are also quasi-isotropic. A close examination of the progressive-damage method shows that its strength properties are not quasi-isotropic

even when the moduli are. In contrast, for the polynomial-invariants method to be developed here, the failure properties will be taken to be quasi-isotropic when the laminate has quasi-isotropic moduli.

The in-plane invariants for quasi-isotropic behavior are

$$(\sigma_{11} + \sigma_{22}) \quad and \quad \left(\sigma_{12}^2 - \sigma_{11}\sigma_{22}\right)$$

Taking the polynomial expansion in these invariants and truncating at second-degree terms gives the failure criterion as

$$\left(\frac{1}{T} - \frac{1}{C}\right)(\sigma_{11} + \sigma_{22}) + \frac{1}{TC}(\sigma_{11} + \sigma_{22})^2 + \frac{1}{S^2}\left(\sigma_{12}^2 - \sigma_{11}\sigma_{22}\right) \leq 1 \tag{12.5}$$

where T and C are the in-plane uniaxial tensile and compressive strengths, and S is the in-plane shear strength.

Many or perhaps most fiber composites applications involve orthotropy. For the case of orthotropy, the full three-dimensional form for polynomial invariants involves twelve independent properties. When specialized to the case of fiber composite laminates, with the in-plane and out-of-plane forms being decomposed, the resultant polynomial-invariants form for the in-plane case involves six properties and is given by

$$\begin{aligned}\left(\frac{1}{T_{11}} - \frac{1}{C_{11}}\right)\sigma_{11} + \frac{1}{T_{11}C_{11}}\sigma_{11}^2 + \left(\frac{1}{T_{22}} - \frac{1}{C_{22}}\right)\sigma_{22} \\ + \frac{1}{T_{22}C_{22}}\sigma_{22}^2 + \lambda_{12}\sigma_{11}\sigma_{22} + \frac{\sigma_{12}^2}{S_{12}^2} \leq 1\end{aligned} \tag{12.6}$$

where the Ts and Cs and S_{12} have the obvious strength identifications, but λ_{12} is yet another independent strength property.

These two expressions are the final results from the polynomial-invariants method for the quasi-isotropic and the orthotropic cases. Three strength properties are required for the former case, and six for the latter. These results are not subject to the fiber-dominated condition used in the preceding development of progressive damage.

For purposes of comparing these results with those of fiber-dominated progressive damage, the numbers of independent strength properties in (12.5) and (12.6) will be reduced somewhat. First, for the quasi-isotropic

case, in the special case where $T = C$, for the polynomial-invariants method to produce the same slope as that of progressive damage, where the envelope crosses the axes then determines S in terms of T as

$$S^2 = \frac{3}{8} T^2 \tag{12.7}$$

Now, in the case where $T \neq C$, the similar procedure gives the best fit between the two methods as

$$S^2 = \frac{3}{8} T C \tag{12.8}$$

In the orthotropic case, the property λ_2 can be determined (not uniquely) in terms of the Ts and Cs as

$$\lambda_{12} = -\frac{2}{3\sqrt{T_{11} C_{11} T_{22} C_{22}}} \tag{12.9}$$

where the 2/3 factor is found by requiring that the form (12.9) reduce to the proper result in the quasi-isotropic specialization of the orthotropic form.

A fiber-dominated estimate of strength property S_{12} in terms of the usual known properties is given by

$$S_{12} = \sqrt{\frac{T_x C_x}{6}} \left[1 - (c_0 + c_{90})\right] \tag{12.10}$$

where T_x and C_x are the lamina-level fiber direction strengths and c_0 and c_{90} are the volume fractions of the 0- and 90-degree lamina in the orthotropic laminate.

With relation (12.8) the properties needed for the quasi-isotropic laminate reduce to only T and C. With relations (12.9) and (12.10) the properties needed for the orthotropic laminate reduces to only T_{11}, C_{11}, T_{22}, and C_{22}. Of course, it would be far preferable to experimentally determine all three properties in the quasi-isotropic case and all six properties in the orthotropic case, but there are situations where this is difficult, and only the uniaxial tests can readily be performed.

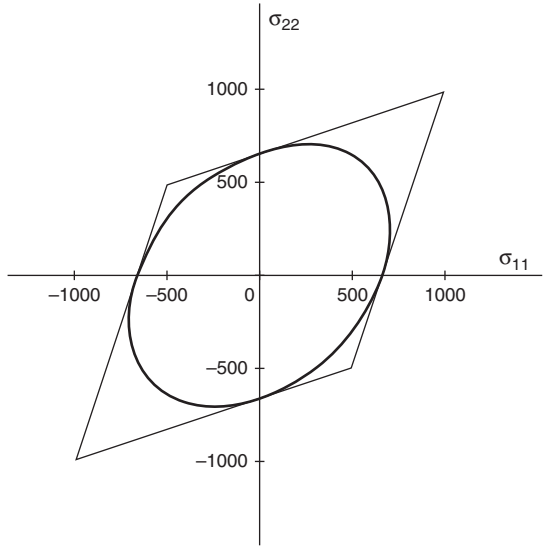


Fig. 12.5 *Quasi-isotropic failure, polynomial invariants, $T = C$.*

Fig. 12.5 shows the polynomial-invariants quasi-isotropic prediction for $T = C$, and Fig. 12.6 for $T \neq C$—both from (12.5) and using (12.7) and (12.8) respectively. Also shown are the predictions from progressive damage, with both methods calibrated to give the same uniaxial strengths in the two directions.

It is now apparent that the polynomial-invariants prediction would provide a more realistic model of the data in Fig. 12.4 than would progressive damage.

Fig. 12.7 shows the orthotropic laminate example failure prediction from polynomial invariants, (12.6) along with (12.9). This orthotropic laminate has the layup pattern specified earlier, (12.4).

The comparisons and differences between polynomial invariants and progressive damage predictions are striking. The polynomial-invariants approach has the rather "smooth" behavior often seen when complicated physical effects along with a great many defects and imperfections are in interaction. Progressive damage appears to be a much more idealized and in effect non-conservative prediction of behavior. In another sense the difference is perhaps in even sharper focus. Progressive damage predicts failure at the lamina level, and presumes that it controls laminate behavior. In contrast, polynomial invariants predicts and is based

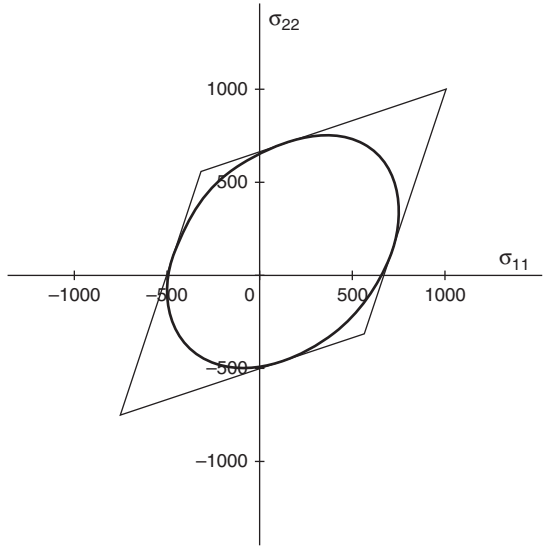


Fig. 12.6 *Quasi-isotropic failure, polynomial invariants,* $T \neq C$.

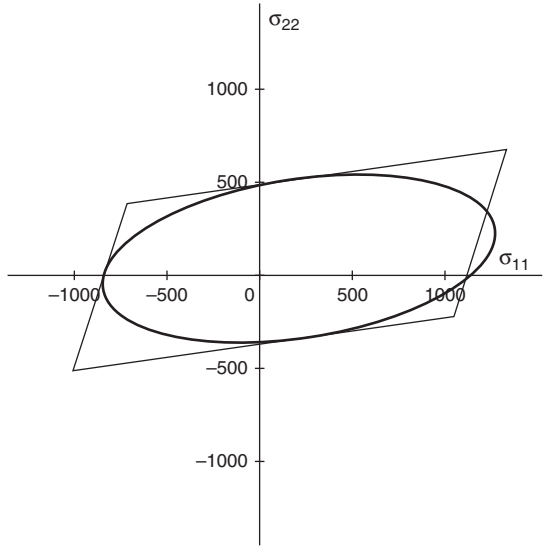


Fig. 12.7 *Orthotropic failure, polynomial invariants.*

only upon failure modes at the laminate scale. Which method is most reflective of physical reality? That remains an open and probably contentious question. The limited data that are available are contradictory concerning this question.

It should be remembered that the progressive damage predictions are based upon the theoretically derived lamina level failure criteria. If one abandons that particular basis and simply adjusts lamina level failure envelopes, then just about any desired laminate level prediction could be obtained through progressive damage. While that might satisfy a particular set of data, there could be no confidence in the generality of that approach. The view favored here is that the polynomial-invariants method has a more solid grounding and is more likely to be generally useful.

There remains the possibility that for preliminary design purposes, progressive damage could be used to determine the uniaxial strengths for any particular laminate configuration, and then that is used in the polynomial-invariants method to produce the complete failure envelope.

A note on terminology may be helpful. The term "polynomial invariants" that is developed and applied here for qausi-isotropic and orthotropic laminates could of course be applied for any symmetry of interest. In fact the isotropic material case of Chapter 4 and subsequent chapters employs polynomial invariants. The three lamina level failure criteria, (12.2), (12.3), and the Tsai–Wu and the Hashin forms, also represent special cases of polynomial invariants.

Up to this point only the in-plane failure characterization has been considered. Now the out-of-plane, delamination, case will be addressed. Using the method of polynomial invariants, it is quickly found that the failure criterion for delamination is given by

$$\left(\frac{1}{T_{33}}-\frac{1}{C_{33}}\right)\sigma_{33}+\frac{1}{T_{33}C_{33}}\sigma_{33}^2+\frac{1}{S_{23}^2}\left(\sigma_{23}^2+\sigma_{31}^2\right)\leq 1 \qquad (12.11)$$

where T_{33} and C_{33} are the thickness direction uniaxial strengths, and S_{23} is the interlaminar shear strength. The form (12.11) is the delamination failure criterion (assumed) for all layups, and the stress components are those transmitted across the interfaces.

Generally, but not necessarily, the failure in the thickness direction is due to delamination between laminae rather than failure within the individual lamina themselves. It is treated as such here. The delamination failure is essentially that of the failure of the extremely thin polymer

layer of locally fluctuating thickness between the lamina. Combinations of through the thickness tension and interlaminar shear are seen to be especially limiting in (12.11).

The failure criterion (12.11) is a fail/no fail type of overall criterion. Sometimes delamination is of a progressing and advancing nature. This type of behavior can be effectively treated using fracture mechanics with a cohesive zone failure model.

In many circumstances the "weak links" with fiber composites are the transverse tensile cracking of the matrix phase and that of delamination. If no matrix cracking is to be allowed then the matrix-controlled failure criterion (12.3) at the lamina level can be used in design to prevent it. Often, however, some small-degree transverse matrix cracking can be tolerated as containable damage. The situation with delamination is usually much more serious, and no delamination at all can be allowed. Any area of delamination destroys the compatibility between the lamina and renders the laminate as disfunctional. Failure criterion (12.11) (or some alternative form) must be used to prevent delamination from occurring under ordinary conditions.

The in-plane and the out-of-plane failure modes have thus far been taken to be independent. In reality there could be some coupling between them. In fact, transverse cracking in the lamina can nucleate delamination in the laminate. It is known that in-plane failure can, in some cases, involve complex combinations of partial delamination with the fiber-controlled failure. The direct method of progressive damage shown here cannot account for such interactive effects. In contrast, the method of polynomial invariants, as applied to the entire laminate, can and does implicitly contain all such effects, including the effects of the lamina stacking sequence.

Further complications that can occur are those of interlaminar edge effects in the testing of laminates. Such effects can obscure the overall objective of determining the intrinsic mechanical properties of failure for the laminate. It is seen that many of the complications that arise with the failure of laminates arise only at the explicit scale of the laminate, and are completely invisible to failure characterization at smaller scales. Progressive damage treats laminate failure as superimposed failure effects from those at smaller scales. Failure is of a completely non-linear character and must include all effects operative up to and including the scale of interest.

It should not be construed that the proper failure criterion for laminates would obviate the need for a proper failure criterion at the lamina level. Both would be equal partners in understanding all failure effects. Indeed, the individual fiber and matrix failure criteria and capability levels are also of great interest as independent matters. Forms for three of these four categories are given in this book, carbon fibers are excepted.

Problem Areas for Study

1. For fiber composites the limiting cases for the directionally varied combinations of lamina are (i) the unidirectional case itself and (ii) the quasi-isotropic combination of lamina. Is it conceivable that a complete understanding of the strength behavior for these two limiting cases could be used to build up a theory for all types of laminates?
2. Delamination in composite material laminates is an especially active and important area. Cohesive zone failure models are often used, especially with numerical modeling. When do these models best apply, and when do regular failure criteria best apply to delamination problems?
3. What is the optimal, or at least the expected, role for the matrix phase in maximizing the strength performance of laminates? Does this favor using polymeric or metallic matrices if strength were the only consideration? Even in the case of using polynomial invariants to describe failure, Section 12.4, it is still important to understand the role of the matrix phase.
4. Fiber composites can be arranged into three-dimensional woven forms, as with textiles. How should an organized approach to the problem of strength for these materials be conducted?
5. Edge effects in laminates always bring in deleterious influences on the strength of laminates. Examine the literature on this subject and propose and assess methods to alleviate the edge condition in the light of the failure criteria presented in this chapter.

References

[12.1] Hinton, M. J., Kaddour, A. S., and Soden, P. D. (2002). "A Comparison of the Predictive Capabilities of Current Failure Theories for Composite Laminates, Judged Against Experimental Evidence," *Composites Sci. and Technology*, **62**, 1725–97.

[12.2] Puck, A. and Schurmann, H. (2002). "Failure Analyses of FRP Laminates by Means of Physically Based Phenomenological Models," *Composites Science and Technology*, **62**, 1633–62.

[12.3] Mayes, S. J. and Hansen, A. C. (2004). "Composite Laminate Failure Analysis using Multicontinuum Theory," *Composites Science and Technology*, **64**, 379–94.

[12.4] Daniel, I. M. (2007). "Failure of Composite Materials," *Strain*, **43**, 4–12.

[12.5] Robbins, D. H. and Reddy, J. N. (2008). "Adaptive Hierarchical Kinematics in Modeling Progressive Damage and Global Failure in Fiber-Reinforced Composite Laminates," *J. Composite Materials*, **42**, 143–72.

[12.6] Tsai, S. W. (Editor) (2008). "Strength and Life of Composites," *J. Composite Materials*, **42**, 1821–988.

[12.7] Tay, T. E., Liu, G., Tan, V. B. C., Sun X. S., and Pham, D. C. (2008). "Progressive Failure Analysis of Composites," *J. Composite Materials*, **42**, 1921–66.

[12.8] Welsh, J. S., Mayes, J. S., and Biskner, A. C. (2007). "Experimental and Numerical Failure Predictions of Biaxially-Loaded Quasi-Isotropic Carbon Composites," *16th Int. Conf. on Composite Materials*, Kyoto, 1–10.

13
Micromechanics Failure Analysis

Taken literally, micromechanics refers to mechanics of materials effects at the 10^{-6} m scale. Usually, however, the term is used more broadly as representing effects at scales somewhat larger than the nanoscale size but much smaller than the macroscopic size. They may center on the micron scale, but they can also reach up and down considerably from that size. Micromechanics certainly offers an appealing look at basic effects at intermediate length scales.

13.1 General Considerations

Many of the problems of special interest in the micromechanics domain are related to inclusions that are dispersed into an otherwise continuous phase. Doing this mixing at or near the micron scale affords an enormous variety of materials combinations that can yield a huge array of enhanced properties. When done properly the combinations can achieve the best attributes of the constituents. When done casually or carelessly it almost always degenerates to the worst from each or all.

The properties that are to be improved include stiffness, strength, toughness, processability, stability, durability, and a host of other, sometimes specialized, properties. Of particular interest here are the cases of improving or optimizing stiffness and strength. Stiffness will be immediately and briefly considered, and then strength will occupy the major effort in this section.

There are many books and countless papers focused upon micromechanics. The vast majority of these deal with the effective stiffness properties. The field of composite materials is the recognized marketplace for such works. For example, *Mechanics of Composite Materials*, Christensen [13.1], is largely given over to the micromechanics analysis of particulate and fibrous reinforcement of a continuous matrix phase.

The objective in most books is to determine the effective stiffness properties of the composite material in terms of the corresponding properties of the phases. There is a very good reason why the effort has been devoted mainly to effective stiffness properties. These stiffness-type problems can be, and are, carefully posed so as to be amenable to complete mechanics-type analysis. The micromechanics situation with strength is much more difficult. This requires some explanation.

The effective stiffness properties involve volume averages over a representative size or element. Thus this procedure is an integration type of result with a consequent smoothing effect. In contrast, failure is usually a point-specific effect located at a defect or a stress concentration point. Even at a stress concentration point there still are defects that are operative at much smaller scales at the same point. Macroscopic strength properties are usually due to a distribution of subscale defects. For these reasons, strength has always been an exceptionally difficult topic.

All of this relates to micromechanics in the following way. If the micromechanics scale embeds the controlling defect then it is well suited for the strength studies. To express this another way, if the prime defect is best and most clearly characterized at the micron level, then micromechanics is the logical choice for proceeding. Otherwise, the advantages of micromechanics for strength studies are far less clear. To go further with this explanation a particular class of materials is needed to gain a firm grasp of the trade offs that must be considered.

Carbon fiber–polymeric matrix composites provide the logical high-performance, high-value example. Carbon fibers (filaments) are about 5–10 microns in diameter, while the separation between them is about an order of magnitude less in dimension. Thus carbon-fiber composites may seem like the perfect fit for micromechanics, but it turns out to be more complicated than it first appears.

The ideal strength approach would involve a micromechanics model that contains all the realistic features such as fiber misalignment, twisted yarns, broken fibers, imperfect fiber bonds with the polymeric matrix, vacuoles in the polymer, resin-rich micro-volumes, and many other non-ideal features. Also there must be fidelity in representing the properties of the two phases: namely, transverse isotropy for the carbon fibers, and appropriate but imperfect properties for the matrix phase. No present computational model can capture all these important details.

To simplify the model there are two options. The first and most obvious approach would be to straighten out the fibers and make them perfectly uniform, with highly idealized properties and embedded in the completely uniform matrix phase. In net effect this simply removes all the defects and flaws that nucleate failure. It is a legitimate micromechanics approach, but even it is still very complex and generally requires a numerical approach. Also, it generally requires parameters, either explicit or hidden, that must be calibrated to macroscopic failure behavior.

The second approach is simpler in one way and more involved in another. It is that of constructing the failure criterion at the macroscopic scale. In this way all of the flaws and defects will automatically be brought in through the macroscopic failure properties that calibrate the theory. This does not mean that micromechanics cannot be helpful in this process. As will be seen, micromechanics can be very helpful, but in a supplementary and elucidating manner. Some of the examples to be presented will show how micromechanics can bring special insights, subtle interpretations, and even pivotal results to the macroscopic theory.

The first approach uses numerical micromechanics to completely and totally construct the failure condition. The second approach uses macroscopic theory to find the appropriate forms for the failure criteria, but then employs micromechanics for special, related conditions. There are many examples of both approaches. An example of the first approach is that of Ha, Huang, Han, and Jin [13.2]. The failure criteria in Chapter 4 of this book, along with work to be presented in this chapter, provides examples of the second approach.

Three micromechanics problems will be examined in this section. The first problem involves one of the properties for matrix-controlled failure in aligned fiber composites. This particular property is particularly troublesome. The second problem relates to particulate inclusions in a continuous matrix phase, which is common practice in product applications and is very important. Then, returning to fiber composites, the third problem is that of load redistribution around broken fibers in unidirectional fiber composites. All three analyses are at the micromechanics level and provide valuable information that is needed in treating the failure of materials at the macroscopic level using macroscopic failure criteria.

13.2 Transverse Shear Strength for Aligned Fiber Composites

At the lamina level in fiber composite laminates the fibers are nominally aligned. It is only at the micromechanics level that one sees the separate phases and the disorder in fiber arrangements. The macroscopic level failure criteria at the lamina level were derived in Chapter 11, and will be recalled below in order to focus upon one of the failure properties that is particularly difficult to characterize. Then the same problem will be examined at the micromechanics scale.

From Sections 11.2 and 11.3 the failure criteria for aligned fiber systems are given by:
Matrix-Controlled Failure

$$\left(\frac{1}{T_{22}} - \frac{1}{C_{22}}\right)(\sigma_{22} + \sigma_{33}) + \frac{1}{T_{22}C_{22}}(\sigma_{22} + \sigma_{33})^2 + \frac{1}{S_{23}^2}(\sigma_{23}^2 - \sigma_{22}\sigma_{33}) + \frac{1}{S_{12}^2}(\sigma_{12}^2 + \sigma_{31}^2) \leq 1 \tag{13.1}$$

Fiber-Controlled Failure

$$-C_{11} \leq \sigma_{11} \leq T_{11} \tag{13.2}$$

In biaxial stress states when $\sigma_{22} = \sigma_{33} = \sigma$ the coefficient of the quadratic term in (13.1) is

$$\left(\frac{4}{T_{22}C_{22}} - \frac{1}{S_{23}^2}\right)\sigma^2 \tag{13.3}$$

For real roots in (13.1) the coefficient in (13.3) must be non-negative. In examining typical sets of failure data for carbon-fiber composites the coefficient in (13.3) is found to be positive in some cases and negative in others. The difference is crucial. If this coefficient is positive then the failure envelope for biaxial stress is an ellipse. But when the coefficient (13.3) goes to zero, the ellipse goes over to an open-ended parabolic form, and the equal biaxial compressive strength becomes infinite. Thus the failure envelope is extremely sensitive to the sign and size of the coefficient in (13.3), which itself involves the small difference between

nearly the same numbers. This sensitivity is so great that it places nearly impossible demands upon the experimental accuracy of the failure data for T_{22}, C_{22}, and S_{23}. To confront this problem it is advantageous to turn to micromechanics to see how it can supply the guidance and enlightenment needed here.

From a macroscopic, mathematical-symmetry point of view, all three properties, T_{22}, C_{22}, and S_{23} are completely independent. However, from a physical point of view for fiber composites there is a reasonable possibility that T_{22}, C_{22}, and S_{23} are related in some manner. The situation is like that in the case of isotropy where T, C, and S are independent properties, but S can be expressed in terms of T and C when a certain eminently reasonable physical condition is invoked. A similarly useful relationship is sought here for S_{23}, giving it in terms of T_{22} and C_{22}. It develops as follows.

All three properties, T_{22}, C_{22}, and S_{23} occur in the matrix-controlled and dominated failure criterion (13.1). Accordingly, the focus should be placed upon the behavior of the matrix phase in carbon-fiber composites under transverse loading. The failure criterion for an isotropic matrix material from several of the previous sections is

$$\left(1 - \frac{T}{C}\right)\hat{\sigma}_{ii} + \frac{1}{2}\left[(\hat{\sigma}_1 - \hat{\sigma}_2)^2 + (\hat{\sigma}_2 - \hat{\sigma}_3)^2 + (\hat{\sigma}_3 - \hat{\sigma}_1)^2\right] \leq \frac{T}{C} \qquad (13.4)$$

where the stresses are non-dimensionalized by C and σ_i are the principal stresses. In this case of isotropy the shear strength S is found from (13.4) to be given by

$$S^2 = \frac{TC}{3}$$

Now consider the expected form for S_{23} in the case of fiber composites. This material is taken as transversely isotropic, which in the 2–3 plane simply has planar isotropy. Following the form for three-dimensional isotropy, the proper form for planar isotropy is taken to be

$$S_{23}^2 = \beta T_{22} C_{22} \qquad (13.5)$$

where β has some specific non-dimensional value that is to be determined.

In order to determine β in (13.5), micromechanics will be employed. A particular case will be examined with specific properties to find β. As the set of guiding typical properties for an epoxy-resin matrix and carbon-fiber composite take:
Calibrating Case

$$\left.\begin{aligned} \nu &= \frac{1}{3} \\ \frac{T}{C} &= \frac{2}{3} \end{aligned}\right\} \textit{Matrix}$$

$$\left.\begin{aligned} E_{11} &>> E_{22} \\ \frac{T_{22}}{C_{22}} &= \frac{1}{3} \end{aligned}\right\} \textit{Composite} \tag{13.6}$$

Because the carbon fibers are much stiffer than the polymeric matrix phase, the transverse deformation of the matrix phase itself will be taken to be that of plane strain deformation. At the micromechanics level an element of the matrix phase itself will be taken to be isolated and subjected to transverse hydrostatic (tensile and compressive) stress states as shown in Fig. 13.1.

The plane strain condition gives $\sigma_{11} = \nu(\sigma_{22} + \sigma_{33})$ or under the two-dimensional hydrostatic stresses

$$\sigma_{11} = 2\nu\sigma$$
$$\sigma_{22} = \sigma_{33} = \sigma$$

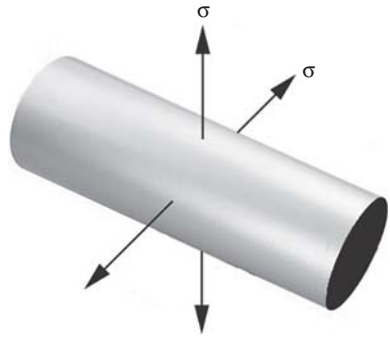


Fig. 13.1 *Matrix in plane strain under two-dimensional hydrostatic stress.*

Using this in (13.4) then gives the ratio of the tensile-to-compressive roots as

$$\Theta = \frac{\sigma^+}{|\sigma^-|} = \frac{-(1+\nu)(1-\lambda) + \sqrt{(1+\nu)^2(1-\lambda)^2 + (1-2\nu)^2\lambda}}{(1+\nu)(1-\lambda) + \sqrt{(1+\nu)^2(1-\lambda)^2 + (1-2\nu)^2\lambda}} \tag{13.7}$$

where ν is the matrix material Poisson's ratio and

$$\lambda = \frac{T}{C}$$

for the isotropic matrix material.

Next, in the composite material the matrix-controlled failure criterion (13.1) will be taken to have the same ratio of transverse hydrostatic tensile and compressive failure stresses as that for the matrix material itself, since the matrix material controls the failure in this situation. The reason for using the equal biaxial stress state rather than a uniaxial stress state is that macroscopic two-dimensional pressure produces a much simpler micromechanics level stress state than does macro uniaxial stress in the transverse direction.

The equal biaxial transverse stresses $\sigma_{22} = \sigma_{33}$ in (13.1) gives the roots as

$$\frac{\sigma_{22}}{C_{22}} = \frac{1}{\Lambda}\left[-(1-\lambda_{22}) \pm \sqrt{(1-\lambda_{22})^2 + \lambda_{22}\Lambda}\right] \tag{13.8}$$

where

$$\lambda_{22} = \frac{T_{22}}{C_{22}}$$

and

$$\Lambda = \left(4 - \frac{T_{22}C_{22}}{S_{23}^2}\right)$$

The equal biaxial stress roots ratio can be formed from (13.8) as

$$\frac{\sigma_{22}^{+}}{\left|\sigma_{22}^{-}\right|}$$

As already mentioned, this composite material ratio from (13.8) will be set equal to the corresponding matrix material ratio from (13.7). This equality of these two ratios at the two different scales is not expected to be true in all fiber composite materials cases, but it will be used here in the special case, (13.6), which is taken to be the guiding case with which to calibrate general behavior and thereby evaluate β in (13.5). That is, the properties in (13.6) are the most typical and common values usually reported, and they will be used here to find the value of β.

Setting $\sigma_{22}^{+}/|\sigma_{22}^{-}|$ from (13.8) equal to the like ratio (13.7) for the matrix material gives an equation to be solved for S_{23}—the property that is so difficult to determine experimentally. After lengthy reduction, but no approximations, it is found that this procedure gives

$$S_{23}^{2} = \left[\frac{(1-\Theta)^{2}}{4\left(1-\frac{T_{22}}{C_{22}}\Theta\right)\left(1-\frac{C_{22}}{T_{22}}\Theta\right)}\right] T_{22}C_{22} \tag{13.9}$$

where Θ is given by the matrix properties in (13.7).

The micromechanics result (13.9) is close to the expected form (13.5), but it is not quite there yet. To bring (13.9) into alignment with (13.5) the non-dimensional coefficient of $T_{22}C_{22}$ in (13.9) will be evaluated for the calibrating special properties in (13.6).

To carry out this process, first use the matrix properties in (13.6) to determine Θ in (13.7). This gives

$$\Theta = \frac{\sqrt{22}-4}{\sqrt{22}+4}$$

Using this value for Θ and T_{22}/C_{22} from (13.6) in (13.9) gives, after much consolidation of terms but no approximations,

$$S_{23}^{2} = \frac{2}{7}T_{22}C_{22} \tag{13.10}$$

This now has the expected form—that of (13.5).

Relation (13.10) is the explicit macroscopic result for the S_{23} property obtained from the micromechanics derivation for carbon-epoxy composites. It has considerable utility in applications, since it obviates the need to experimentally determine S_{23} with all its attendant problems.

With relation (13.10) combined with failure criteria (13.1) and (13.2) there are now five independent failure properties for aligned fiber composites. This is the same as the number of independent elastic properties for transverse isotropy. A further interpretation of (13.10) will now be presented.

It is interesting to observe that the result (13.10) falls within some likely bounds for S_{23} for carbon-fiber composite materials. Consider the general case of aligned fiber composite materials when the fiber stiffness is varied over the full range possible. At the upper limit where the fibers have the same stiffness as the matrix material, then the value for S_{23} must be the same as that for an isotropic material: namely,

$$S_{23}^2 = \frac{T_{22}\, C_{22}}{3}$$

At the lower limit, the foregoing failure forms (13.1) and (13.3) show that there must be

$$S_{23}^2 > \frac{T_{22}\, C_{22}}{4}$$

If S_{23} varies monotonically between these two limits, as it probably does, then it is seen that the result (13.10) with the coefficient of 2/7 fits between the above two limits of 2/6 and 2/8.

Finally, three examples will be presented. These examples have the T_{22}/C_{22} ratios of 1/4, 1/3, and 1/2.4. These cover the usual range found for this properties ratio.

For Example 1 take

$$T_{22} = 50\ MPa$$
$$C_{22} = 200\ MPa$$

Then (13.10) gives S_{23} as

$$S_{23} = 53.5\ MPa$$

In equal biaxial tension and compression, the roots from (13.8) are

$$\sigma_{22}^{+} = 31.6\ MPa$$
$$\sigma_{22}^{-} = -632\ MPa$$

For Example 2 take

$$T_{22} = 50\ MPa$$
$$C_{22} = 150\ MPa$$

This example corresponds to that of the calibrating case in (13.6). From (13.10) there follows

$$S_{23} = 46.3\ MPa$$

and from (13.8) the equal biaxial failure stresses are

$$\sigma_{22}^{+} = 34.6\ MPa$$
$$\sigma_{22}^{-} = -435\ MPa$$

For Example 3 take

$$T_{22} = 50\ MPa$$
$$C_{22} = 120\ MPa$$

From (13.10) there is

$$S_{23} = 41.4\ MPa$$

and from (13.8)

$$\sigma_{22}^{+} = 37.8\ MPa$$
$$\sigma_{22}^{-} = -318\ MPa$$

It is seen that quite large stress levels are required to fail these materials in equal biaxial compressive stress states. Nevertheless these materials can

and must fail in these transverse stress states, as can be reasoned independently. The final form (13.10) derived from micromechanics necessarily must be used with all applications of failure criterion (11.7), (13.1).

13.3 Spherical Inclusion in an Infinite Elastic Medium

Although two of the three micromechanics examples are concerned with fiber composites, particulate inclusion composites are equally important and even much more common. The problem to be examined here is that of a particulate spherical inclusion in an infinite medium. This symbolizes the dilute suspension case. Of course, the spherical inclusion can be of any size, but the vast majority of particulate inclusions used in materials synthesis are within the usual range of micromechanics, taken in the broad sense.

The explicit problem is that of a rigid spherical inclusion in an infinite elastic medium under far-field uniaxial stress. First the elastic solution for the problem will be given, then the strength problem will be examined. This problem was briefly stated in Chapter 10 on fracture. Here the full solution will be given. Although this fundamental elasticity problem surely must be available somewhere in the classical literature, it is not known to the author where that may be. Accordingly, the full-field elastic solution is developed here for use in the correspondingly basic particulate strength example.

The problem is shown in Fig. 13.2. The solution is found from the three coupled, partial differential equations of equilibrium expressed in terms of displacements for an elastic medium containing a rigid spherical inclusion.

The solution is given in spherical coordinates, with the polar axis $\theta = 0$ being in the uniaxial stress direction. The displacement solution is given by

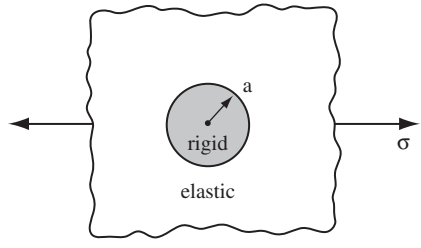


Fig. 13.2 *Rigid spherical inclusion in an infinite elastic medium.*

$$\frac{E}{\sigma}u_r = \frac{(1-2\nu)}{3}\left(r - \frac{a^3}{r^2}\right) + \left(\frac{1+\nu}{3}\right)\left[r - \frac{5\,(5-4\nu)}{4\,(4-5\nu)}\frac{a^3}{r^2} + \frac{9}{4\,(4-5\nu)}\frac{a^5}{r^4}\right]\left(2\cos^2\theta - \sin^2\theta\right)$$

$$\frac{E}{\sigma}u_\theta = (1+\nu)\left[-r + \frac{5\,(1-2\nu)}{2\,(4-5\nu)}\frac{a^3}{r^2} + \frac{3}{2\,(4-5\nu)}\frac{a^5}{r^4}\right]\sin\theta\cos\theta \qquad (13.11)$$

The displacement component u_ϕ vanishes, since the problem is axisymmetric. The symbol σ denotes the applied far-field uniaxial stress.

The complete stress field is found to be given by

$$\begin{aligned}
\frac{\sigma_{rr}}{\sigma} &= \cos^2\theta + \frac{2}{3}\left(\frac{1-2\nu}{1+\nu}\right)\left(\frac{a}{r}\right)^3 \\
&\quad + \frac{1}{4-5\nu}\left[-\frac{5}{6}\,(5-\nu)\left(\frac{a}{r}\right)^3 + 3\left(\frac{a}{r}\right)^5\right]\sin^2\theta \\
&\quad + \frac{1}{4-5\nu}\left[\frac{5}{3}\,(5-\nu)\left(\frac{a}{r}\right)^3 - 6\left(\frac{a}{r}\right)^5\right]\cos^2\theta \\
\frac{\sigma_{\theta\theta}}{\sigma} &= \sin^2\theta - \frac{1}{3}\left(\frac{1-2\nu}{1+\nu}\right)\left(\frac{a}{r}\right)^3 \\
&\quad - \frac{1}{4-5\nu}\left[\frac{5}{12}\,(1-2\nu)\left(\frac{a}{r}\right)^3 + \frac{9}{4}\left(\frac{a}{r}\right)^5\right]\sin^2\theta \\
&\quad + \frac{1}{4-5\nu}\left[-\frac{5}{3}\,(1-2\nu)\left(\frac{a}{r}\right)^3 + 3\left(\frac{a}{r}\right)^5\right]\cos^2\theta \\
\frac{\sigma_{\phi\phi}}{\sigma} &= -\frac{1}{3}\left(\frac{1-2\nu}{1+\nu}\right)\left(\frac{a}{r}\right)^3 \\
&\quad + \frac{1}{4-5\nu}\left[\frac{25}{12}\,(1-2\nu)\left(\frac{a}{r}\right)^3 - \frac{3}{4}\left(\frac{a}{r}\right)^5\right]\sin^2\theta \\
&\quad + \frac{1}{4-5\nu}\left[-\frac{5}{3}\,(1-2\nu)\left(\frac{a}{r}\right)^3 + 3\left(\frac{a}{r}\right)^5\right]\cos^2\theta \\
\frac{\sigma_{r\theta}}{\sigma} &= \left[-1 + \frac{5}{2}\left(\frac{1+\nu}{4-5\nu}\right)\left(\frac{a}{r}\right)^3 - \frac{6}{(4-5\nu)}\left(\frac{a}{r}\right)^5\right]\sin\theta\cos\theta
\end{aligned} \qquad (13.12)$$

The other two shear stress components vanish.

At $r = a$, next to the bonded surface of the rigid spherical inclusion, the stresses in the elastic matrix medium become

$$\begin{aligned}
\sigma_{rr} &= \frac{3(1-\nu)}{(4-5\nu)}\left[\frac{3}{1+\nu} - \frac{5}{2}\sin^2\theta\right]\sigma \\
\sigma_{\theta\theta} = \sigma_{\phi\phi} &= \frac{3\nu}{(4-5\nu)}\left[\frac{3}{1+\nu} - \frac{5}{2}\sin^2\theta\right]\sigma \\
\sigma_{r\theta} &= \frac{-15(1-\nu)}{2(4-5\nu)}\sigma\sin\theta\cos\theta
\end{aligned} \tag{13.13}$$

Two failure examples will be considered. In both of these examples the critical stress state is found to be at $r/a = 1$ and $\theta = 0$. This is where the tensile stresses dominate for the brittle materials to be considered. The stresses at this point are

$$\begin{aligned}
\sigma_{rr} &= \frac{9(1-\nu)\sigma}{(1+\nu)(4-5\nu)} \\
\sigma_{\theta\theta} = \sigma_{\phi\phi} &= \frac{9\nu\sigma}{(1+\nu)(4-5\nu)}
\end{aligned} \tag{13.14}$$

The two examples to be considered have the same Poisson's ratio and different values of T/C, given by

$$\begin{aligned}
\nu &= \frac{1}{3} \\
\frac{T}{C} &= \frac{1}{2} \text{ and } \frac{1}{4}
\end{aligned}$$

The $T/C = 1/2$ example is that for the range of fairly brittle polymers, and the $T/C = 1/4$ case would be for ideal ceramic-type materials. Many common ceramics would have T/C values even considerably less than $T/C = 1/4$.

The failure criterion is given by

$$\left(1-\frac{T}{C}\right)\hat{\sigma}_{ii}+\frac{1}{2}\left[(\hat{\sigma}_1-\hat{\sigma}_2)^2+(\hat{\sigma}_2-\hat{\sigma}_3)^2+(\hat{\sigma}_3-\hat{\sigma}_1)^2\right]\leq\frac{T}{C} \quad (13.15)$$

The fracture mode criterion is

$$\begin{aligned}\hat{\sigma}_1 &\leq \frac{T}{C}\\ \hat{\sigma}_2 &\leq \frac{T}{C} \qquad for \quad \frac{T}{C}\leq\frac{1}{2}\\ \hat{\sigma}_3 &\leq \frac{T}{C}\end{aligned} \quad (13.16)$$

The stresses are normalized by the uniaxial compressive strength C.

From (13.14) the stresses at $r/a = 1$ and $\theta = 0$ in both examples are

$$\begin{aligned}\sigma_{rr} &= \frac{27}{14}\sigma\\ \sigma_{\theta\theta} &= \sigma_{\phi\phi} = \frac{27}{28}\sigma\end{aligned}$$

Using the above stresses in the failure criterion (13.15) gives

$$\left(\frac{27}{28}\right)^2\hat{\sigma}^2+\frac{27}{7}\left(1-\frac{T}{C}\right)\hat{\sigma}-\frac{T}{C}=0 \quad (13.17)$$

It is the failure criterion (13.15), not the fracture criterion (13.16), that is found to be critical in both examples.

For $T/C = 1/2$ the failure criterion (13.17) becomes

$$\hat{\sigma}^2+\frac{56}{27}\hat{\sigma}-2\left(\frac{14}{27}\right)^2=0$$

The solution for the far-field stress, σ, at failure is then found to be

$$\begin{aligned}\sigma &= \frac{56}{27}\left(\sqrt{\frac{3}{2}}-1\right)T\\ &= 0.466\,T\end{aligned}$$

For the $T/C = 1/4$ case the failure criterion (13.17) gives

$$\hat{\sigma}^2 + \frac{28}{9}\hat{\sigma} - \left(\frac{14}{27}\right)^2 = 0$$

The solution for the far-field stress at failure is then

$$\begin{aligned}\sigma &= \frac{56}{27}\left(\sqrt{10} - 3\right) T \\ &= 0.337\,T\end{aligned}$$

This case is seen to have the proverbial stress concentration factor of 3. The first example has a stress concentration factor between 2 and 3 but closer to 2. The term "stress concentration factor" is used here as emanating from the failure criterion involving the full stress state, not that of any single component of stress, which would be meaningless when there are multiple components.

The failure examples just considered are point-specific. That is, the location is found where incipient failure first occurs. Does this imply macroscopic failure, or would the local failure simply remain as local damage? There is no general, all-inclusive answer to this question, and each problem must be considered individually.

If the material were a very ductile metal, then the point-specific failure would probably remain as a local plastic flow region. In the problem of a spherical inclusion in a very ductile medium the first plastic flow would occur at $r = a$, and probably near where the shear stress is a maximum at $\theta = \pi/4$.

In the two examples given above, the first failure occurs at the pole, $\theta = 0$. In both examples the local failures would probably lead to macroscopic failures. Especially in the second example for $T/C = 1/4$, and with the three-dimensional far-field tensile stress state, it is very probable that the local failure would take the form of a brittle crack(s) that would easily and quickly propagate throughout the material.

In treating problems of this type using isotropic macroscopic failure criteria, micromechanics analysis such as given here are very helpful in understanding what nucleates the macroscopic failure.

13.4 Load Redistribution in Aligned Fiber Composites

The function of the matrix phase in aligned fiber composite materials is often misunderstood and usually completely undervalued. The matrix is seen as little more than a space-filling medium between the fibers that can accommodate fittings and attachments. The real and vital function of the matrix phase is best understood through micromechanics. One example has already been given in the second section of this chapter, and another important function will now be examined.

In the introduction to this chapter it was mentioned that the state of flaws and defects are usually the nucleating sites for failure in fiber composites. Chief among such defects are broken filaments in the fiber bundles, fiber slack, and variability along the fiber filaments. The matrix phase supplies the palliative that largely overcomes these defects. Without the matrix phase it would be virtually impossible to utilize the extraordinary properties that are inherent in carbon fibers and other high-performance fibers.

The load rebalancing or redistribution function of the matrix phase in transferring load around broken fibers provides the perfect example to illuminate the true role of the matrix phase. As will be seen, the key to performance in composites lies in the interaction between the high-performance fibers and the coordinating matrix phase. This is an interaction that must extract the most from the fibers but cannot do so without the perfect complement of matrix properties.

13.4.1 Boundary Layer Theory

It is advantageous to start with a brief summary and example of the macroscopic boundary layer theory for highly anisotropic fiber composites, because it will pinpoint the properties and features that are needed in the micromechanics model for load redistribution. This boundary layer theory for fiber composites was developed by Pipkin and by Spencer and others. The outline given here follows from that of Christensen [13.1], which in turn follows their works.

The boundary layer theory starts by recognizing that at the lamina level carbon fiber composites are highly anisotropic. Accordingly, a simplified method of analysis is sought for this class of materials. It begins by identifying a small non-dimensional parameter that can be used in the

analysis. This rather obvious small parameter is given by

$$\varepsilon^2 = \frac{\mu_{12}}{E_{11}/(1 - \nu_{12}\nu_{21})} \tag{13.18}$$

where the lamina is in a plane stress state, with E_{11} the fiber direction modulus, μ_{12} the axial shear modulus, and the other terms the Poisson's ratios. All other moduli-type properties are taken to be of the same order as μ_{12}. One more step is needed. In order to provide a clear view of the thin boundary layer, the transverse direction coordinate x_2 is "stretched" to the coordinate η by taking

$$x_2 = \varepsilon\eta \tag{13.19}$$

The complete formulation can be arranged with terms expressed as powers of ε. By retaining only the lowest-order terms in ε, a fortunate thing happens. The entire formulation is found to reduce to and be governed by the following differential equations for the displacements:

$$\begin{aligned} \frac{\partial^2 u_1}{\partial x_1^2} + \frac{\partial^2 u_1}{\partial \eta^2} &= 0 \\ \frac{\partial^2 u_2}{\partial \eta^2} &= 0 \end{aligned} \tag{13.20}$$

Thus the equilibrium formulation is given by solving Laplace's equation rather than the much more complicated biharmonic equation.

The problem to be solved here is the basic Green's function problem of a concentrated force applied within the full plane, Fig. 13.3. The force P is the force per unit thickness, acting in the fiber direction.

The displacements are solved from (13.20), and the stresses, to consistent order in ε, are then given by

$$\begin{aligned} \sigma_{11} &= -\frac{P}{2\pi\varepsilon}\,\frac{x_1}{(x_1^2 + \eta^2)} \\ \sigma_{12} &= -\frac{P}{2\pi}\,\frac{\eta}{(x_1^2 + \eta^2)} \\ \sigma_{22} &= 0 \end{aligned} \tag{13.21}$$

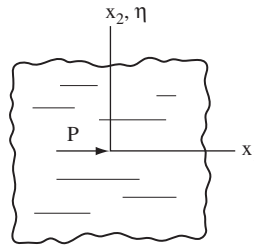


Fig. 13.3 *Concentrated force in an infinite plane.*

For high anisotropy, as occurs here, the Poisson's ratios in (13.18) can be neglected, and the stresses (13.21) at $x_2 = 0$ and at $x_1 = 0$ are given by

$$\begin{aligned} \sigma_{11}\big|_{x_2=0} &= -\frac{P}{2\pi}\sqrt{\frac{E_{11}}{\mu_{12}}}\left(\frac{1}{x_1}\right) \\ \sigma_{12}\big|_{x_1=0} &= -\frac{P}{2\pi}\sqrt{\frac{\mu_{12}}{E_{11}}}\left(\frac{1}{x_2}\right) \end{aligned} \tag{13.22}$$

It is seen from these results that the characterizing stresses are σ_{11} and σ_{12}. Both have singularities, but σ_{11} decays much more slowly with distance x_1 from P than σ_{12} decays with respect to distance x_2 from P. Stress σ_{22} does not even enter the problem. All of this is the boundary layer effect. For infinitely stiff fibers the load is transmitted indefinitely along a singular line with no load being diffused out of it, since the lateral shear stress vanishes.

Only two properties are involved in the entire analysis. These are the modulus E_{11} controlled by the fibers, and μ_{12} controlled by the matrix. All of this enough to provide valuable and complete guidance in the micromechanics problem of interest, which is discussed next.

13.4.2 Load Redistribution Problem

The simplest form of the load redistribution around a broken fiber is as shown in Fig. 13.4.

This two-dimensional micromechanics problem has one broken fiber surrounded by two continuous ones. The two surrounding fibers will have an overload condition at $x = 0$, but the overload will gradually diffuse

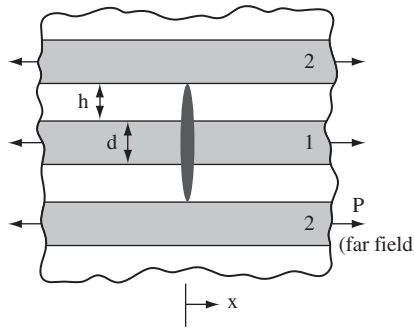


Fig. 13.4 *Broken fiber stress redistribution problem.*

into the broken fiber at positions way from the break. Next is where the boundary layer behavior shown above comes into use. The boundary layer solution shows that the load in the x direction is totally controlled by the fibers, and that the load diffuses from one fiber to another purely through the shear deformation of the matrix. In addition, only the related two properties are involved.

Therefore, the entire micromechanics problem is controlled by axial load in the fibers and shear stress in the matrix. This condition is commonly known as "shear-lag" and is usually used on an *ad hoc* basis with no specific justification. The approach here has a rigorous justification from boundary layer theory, even though the equilibrium equations are violated in the matrix phase and only satisfied in the fiber phase as resultants. This provides the micromechanics analysis with a much more firm foundation than it would otherwise have, and it is understood why this can only be done when the fibers are much stiffer than the matrix phase. Although the shear-lag approach has been around for a very long time, the first rational treatment of it with all its implications for fiber composites was only recently presented by Xia, Curtin, and Okabe [13.3].

With reference to the model in Fig. 13.4, then for fiber equilibrium including shear stresses transmitted through the matrix, the fiber displacements and loads are found to be

$$\begin{aligned} u_1 &= \frac{P}{E_f}\left[\frac{x}{d} + \frac{1}{2\lambda}\left(1 + 2e^{-\lambda\frac{x}{d}}\right)\right] \\ u_2 &= \frac{P}{E_f}\left[\frac{x}{d} + \frac{1}{2\lambda}\left(1 - e^{-\lambda\frac{x}{d}}\right)\right] \end{aligned} \tag{13.23}$$

and

$$\begin{aligned} P_1 &= P\left(1 - e^{-\lambda\frac{x}{d}}\right) \\ P_2 &= P\left(1 + \frac{1}{2}e^{-\lambda\frac{x}{d}}\right) \end{aligned} \tag{13.24}$$

where

$$\lambda = \sqrt{\frac{3\mu_m d}{E_f h}} \tag{13.25}$$

and where P is the far-field load in the fibers.

The shear stress in the matrix is given by

$$\tau = \frac{\mu_m}{h}\left(u_2 - u_1\right) \tag{13.26}$$

The overall far-field stress in the composite from Fig. 13.4 is specified by

$$\sigma = \frac{P}{h + d} \tag{13.27}$$

With these results the maximum stress in the fibers and the maximum shear stress in the matrix are both found to occur at $x = 0$ and are given by

$$\begin{aligned} \sigma_f &= \frac{3}{2}\left(1 + \frac{h}{d}\right)\sigma \\ \tau_m &= \frac{\sqrt{3}}{2}\left[\left(1 + \frac{h}{d}\right)\sqrt{\frac{d\mu_m}{hE_f}}\right]\sigma \end{aligned} \tag{13.28}$$

Thus the adjacent fiber overload is at the 50% level, as is apparent directly in Fig. 13.4. Not at all obvious by inspection but equally important is the maximum shear stress (13.28) in the matrix needed to transfer the load around the broken fiber.

The ratio of these maximum stresses in the two phases is

$$\frac{\sigma_f}{\tau_m} = \sqrt{\frac{3hE_f}{d\mu_m}} \tag{13.29}$$

Now, as an example take $d/h = 3$ such that the volume fraction of the fiber phase is 3/4, giving (13.29) as

$$\frac{\sigma_f}{\tau_m} = \sqrt{\frac{E_f}{\mu_m}} \tag{13.30}$$

This ratio of the maximum stresses in the two phases is to be compared with the ratio of their corresponding strengths.

For the properties to use with (13.30) take the carbon fibers and epoxy matrix with

$$E_f = 250 \ GPa$$

$$\mu_m = 1.25 \ GPa$$

and for the corresponding strengths take

$$T_f = 5 \ GPa$$

$$S_m = 50 \ MPa$$

So the ratio of maximum stresses and the ratio of the corresponding strengths are

$$\frac{\sigma_f}{\tau_m} = 14.1 \qquad \frac{T_f}{S_m} = 100 \qquad \begin{matrix} f \sim fiber \\ m \sim matrix \end{matrix}$$

It is seen that in this broken fiber stress redistribution problem the matrix fails at a lower macroscopic stress than does the fiber phase. The matrix phase is more critical than the fiber phase, even though the problem is nominally that of load carried only in the fiber direction. Of course, there is some latitude or uncertainty in the values of the above properties and the proper value for the volume fraction of fibers.

This result from the simple three-fiber model of Fig. 13.4 can be checked against much more elaborate models of the same basic load redistribution problem. Chou [13.4] has produced such a solution for a single fiber break in a full field of (two-dimensional) fibers. The basic conclusion just found remains unchanged. Considering the properties uncertainties, it follows that both types of failure can and will occur in the general situation of axially loaded aligned fiber composites.

This micromechanics analysis of the load redistribution problem shows that the macroscopic fiber-controlled failure criterion actually involves a very complex combination of fiber failure and possibly concurrent matrix failure. This provides a powerful incentive for treating failure criteria at the macroscopic scale, because both of the complex effects are implicitly present at the macroscale. Nevertheless, this basic conclusion can only be seen and appreciated through the micromechanics scale failure analysis.

Although it is very difficult to quantitatively predict macroscopic failure from microscale, nanoscale, or atomic scale considerations, these subscale analyses are still vital and indeed irreplaceable. All new materials and most major advancements in existing materials start as concepts at the atomic, nano, or microscale levels. It is extremely important to understand the physical behavior at all scales, even though the applications are dominantly macroscale.

Problem Areas for Study

1. What are the limitations of using micromechanics to describe the macroscopic failure of aligned fiber composites? For example, micromechanics combined with fracture mechanics works well for the transverse matrix cracking between the fibers, but micromechanics is of no help with the fiber kinking under compressive stress conditions in the fiber direction. What about other modes of failure?
2. The problem of an isotropic thin sheet under uniaxial stress and containing a circular hole has a stress concentration factor of 3. There is no need for a separate failure criterion, as the stress concentration factor is the failure criterion! This is because only a single component of stress is active at the critical location. How many ideal strength problems of this type are there in existence?
3. In Section 13.2, micromechanics is used to reduce the number of independent failure properties for a transversely isotropic aligned fiber

composite from six to five. The number five is also the number of independent elastic properties for the same type of material. The same effect occurred with the isotropic materials failure theory discussed in Chapter 4, having two each of stiffness and strength properties. Is this coordination of the numbers of independent properties for elastic stiffness and for strength just a coincidence, or is there a fundamental underlying principle involved? If so, what is its basis?

4. High-impact polystyrene contains a suspension of polystyrene inclusions within an elastomeric phase, which actually is embedded within the continuous phase of polystyrene. Thus this is a composite within a composite. All of this occurs on the microscale. The overall material is thereby toughened against impact failure. What physical mechanism accounts for the improvement in the failure behavior? What other toughening mechanisms could be instituted on the microscale?
5. Section 13.4 presents a two-dimensional analysis of load transfer around a broken fiber in a unidirectional fiber system. Without working out an explicit three-dimensional problem of the same type, reason whether the three-dimensional case would be expected to have substantially the same results as in the two-dimensional case, or be substantially different.

References

[13.1] Christensen, R. M., 2005, *Mechanics of Composite Materials*, Dover, New York.

[13.2] Ha, S. K., Huang, Y., Han, H. H., and Jin, K. K., 2010, "Micromechanics of Failure for Ultimate Strength Predictions of Composite Laminates," *J. Composite Materials*, **44**, 2347–61.

[13.3] Xia, Z., Curtin, W. A., and Okabe, T., 2002, "Green's Function vs. Shear–Lag Models of Damage and Failure in Fiber Composites," *Composites Science and Technology*, **62**, 1279–88.

[13.4] Chou, T. W., 1992, *Microstructural Design of Fiber Composites*, Cambridge University Press, Cambridge.

14
Nanomechanics Failure Analysis

The nanoscale extends to much smaller length scales of physical effects than does the microscale. It approaches the size of atoms and thereby that of the interaction between atoms, and as such it covers all of the common molecular sizes. Effects at the nanoscale are invisible to the influence of the microscale and macroscale flaws and defects that determine so much of the failure behavior at the macroscopic size. Nevertheless, the nanoscale provides the most basic window on materials capability, and thereby provides the template onto which the much larger defect sizes must be overlaid. It is the combination of all these effects at all scales that determines materials performance.

The most common example of nanoscale-influenced behavior is that of dislocations and their flow and interaction. That subject is so extremely well covered and understood that there is no need to enter into it here. One of the most revealing and far-reaching modern discoveries of special nanostructures are those in the domain of the element carbon. This chapter will closely examine the properties of graphene—the sheet form of bonded carbon atoms at a single atomic layer of thickness. This work will be so revealing that it will immediately suggest a theoretical method by which to examine those elements in the Periodic Table that form strong solids. In particular, the ductile-versus-brittle failure properties of the elements will be characterized through a nanoscale analysis.

14.1 Graphene Nanostructure

Carbon is an element of extraordinary properties. The carbon–carbon bond possesses large magnitude cohesive strength through its covalent bonds. Elemental carbon appears in a variety of forms from common graphite to high-technology carbon fibers and on to high-purity diamond, and finally to the modern, ideal nanoscale forms of Fullerenes, nanotubes, and graphene. When Fullerenes (Bucky-Balls) first came into initial exposure and testing, the form was termed by some as “The

Molecule of the Year." Later, of course, it led to multiple Nobel Prizes. Graphene is the perfect planar form of a sheet of carbon atoms of only one atom of thickness. Fullerenes and single-walled nanotubes provide the spherical and cylindrical forms that graphene can take. The planar form—graphene—will be examined here, but it also represents the inherent properties available and exploitable in Fullerenes and nanotubes.

Carbon has the atomic number 6. It has two electrons in an inner shell and four electrons in the outer shell of orbitals. In graphene the four bonding electrons actually form three bonds with nearest-neighbor atoms, with one of these being a double bond. The two-dimensional nature of graphene then has the atoms arranged in an hexagonal pattern, as shown in Fig. 14.1, each with three nearest-neighbors. The atomic centers are at the nodes of the pattern, and the outer shells of electrons meet midway between nodes.

The hexagonal arrangement of the atoms has a six-fold symmetry that assures isotropy of the in-plane, large-scale stiffness and compliance properties. In order to model the average stiffness properties of graphene, the hexagonal pattern of atoms is idealized as an hexagonal pattern of connected elastic members that possess axial and bending stiffness characteristics that will be adjusted to represent the bond-stretching (and contraction) and the bond-bending (distortion) effects. Each individual elastic member between atomic centers is taken with the axial and bending properties specified by the stiffness coefficients

$$\begin{aligned} k_A &= \frac{AE}{l} \\ k_B &= \frac{12EI}{l^3} \end{aligned} \tag{14.1}$$

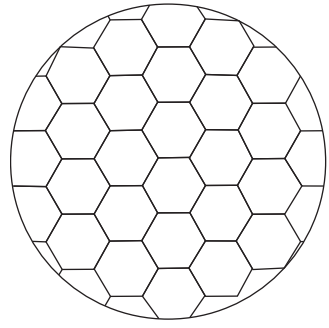


Fig. 14.1 *Hexagonal pattern of graphene atoms.*

where l is the distance between atomic centers and the other properties have the usual mechanics designations. The stiffness coefficient k_B is that for the relative displacements of the beam-member ends with no end rotations. This overall treatment of the mechanical properties for graphene is motivated by the original conception of Bohr that for some purposes individual atoms could be taken as effectively being elastic bodies.

The hexagonal pattern of elastic members is analyzed to obtain the effective elastic properties of the sheet of carbon atoms. The bending behaviors of the connecting elastic members are represented by Bernoulli–Euler beam theory. At the atomic centers (nodes) the displacements and their first derivatives (slopes) of the members must be compatible and continuous, and equilibrium of the forces and moments are required at the nodes (atomic centers). The analysis to ensure the equilibrium conditions is algebraically fairly long, but with no complications or approximations. Thereafter, an homogenization process is applied to average over the complete assemblage of elastic members.

The end results of the analysis are the in-plane elastic properties given by the two-dimensional Poisson's ratio, Young's modulus, shear modulus, and bulk modulus. These are respectively found to be given by the Poisson's ratio form

$$\nu_{2D} = \frac{1-\kappa}{1+3\kappa} \qquad \left(-\frac{1}{3} \le \nu_{2D} \le 1\right) \tag{14.2}$$

and the three moduli

$$\begin{aligned} E_{2D} &= \frac{4\kappa}{\sqrt{3}\,(1+3\kappa)} k_A \\ \mu_{2D} &= \frac{\kappa}{\sqrt{3}\,(1+\kappa)} k_A \\ K_{2D} &= \frac{k_A}{2\sqrt{3}} \end{aligned} \tag{14.3}$$

where

$$\kappa = \frac{k_B}{k_A} \qquad (0 \le \kappa \le \infty) \tag{14.4}$$

These properties are the controlling, in-plane mechanical properties for the graphene-type arrangement of carbon atoms. Since the graphene sheet is only one atom thick there is no characteristic thickness dimension, and

the three moduli in (14.3) have units of force per unit length, not per unit area. The limits on Poisson's ratio in (14.2) are found by letting k_A and k_B in (14.4) assume all non-negative values. The single elastic member between two neighboring atoms is taken as shown in Fig. 14.2.

Appropriate to the two-dimensional nature of the problem for graphene, the elastic members connecting atomic centers are taken as planar forms of length l and width d. The thickness of the elastic members are left as indeterminate and need not be specified here to obtain the variable κ. In this planar case the κ from (14.1) and (14.4) becomes simply

$$\kappa = \left(\frac{d}{l}\right)^2 \tag{14.5}$$

In obtaining (14.5) the factor of 12 in (14.1) is cancelled by a factor of 12 for the moment of inertia, I, for the connecting members.

Physically, it is obvious that the elastic member width d could not possibly extend beyond the outer shell of the electrons, thus

$$0 \leq \frac{d}{l} \leq 1 \tag{14.6}$$

With (14.5), (14.6) now becomes

$$0 \leq \kappa \leq 1 \tag{14.7}$$

Furthermore, with the physical restrictions (14.6) and (14.7) the Poissons's ratio (14.2) now has the limits

$$0 \leq \nu_{2D} \leq 1 \tag{14.8}$$

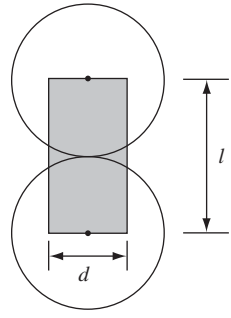


Fig. 14.2 *Idealized elastic member between two neighboring carbon atoms.*

The case of $\nu_{2D} = 0$ is particularly interesting. This occurs at $\kappa = 1$, which means that the bond-stretching resistance is perfectly balanced with the bond-bending resistance.

It is clear why there cannot be negative values for Poisson's ratios in this case of graphene with its covalent bonding. To allow negative Poisson's ratios would require violation of the physical restriction (14.6), which would be fundamentally unrealistic and unacceptable. With this new understanding it is entirely appropriate to say that the positive Poisson's ratios are physically realizable but negative values are not.

The use of the Bernoulli–Euler theory for the bending behavior extends up to an aspect ratio of d/l that would not be used with ordinary structural members. But the use here is totally different, with its providing the complete bond-bending complement to the bond-stretching effect without introducing any additional properties that must be specified independently.

By far the most interesting and most informative property in the group of four properties in (14.2) and (14.3) is the Poisson's ratio. Inverting the form (14.2) to obtain κ in terms of ν_{2D} gives

$$\kappa = \frac{1 - \nu_{2D}}{1 + 3\nu_{2D}} \tag{14.9}$$

Comparing (14.2) and (14.9) it is seen that the relation between ν_{2D} and κ is form-invariant. From (14.2) or (14.9) and (14.7) the Poisson's ratio ν_{2D} and the non-dimensional bond-bending/bond-stretching variable κ ranges over the values

ν_{2D}	κ
0	1
1/9	2/3
1/5	1/2
1/3	1/3
1/2	1/5
2/3	1/9
1	0

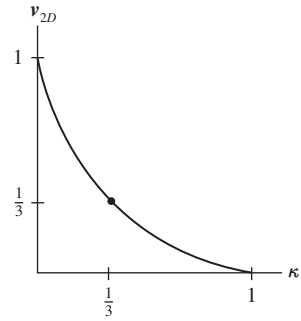


Fig. 14.3 *ν_{2D} versus κ for graphene.*

The graph of ν_{2D} versus κ, (14.2), is shown in Fig. 14.3.

It is seen from (14.2) and (14.9) that the ν_{2D} versus κ is symmetric about a 45-degree axis that goes through the origin and the central point of

$$\nu_{2D} = \kappa = \frac{1}{3}$$

This value for ν_{2D} and κ has a very special significance. It is at this point that the behaviors of ν_{2D} and κ interchange roles with each other. This is a special turning point or transition point.

There is very probably a change of physical behavior signified by $\nu_{2D} = 1/3$. The terms "ductile" and "brittle" could perhps be used to describe the change. With this "perfect" material—graphene—perhaps the best applicable descriptions would refer to extensibility up to failure such that

$$For \quad \nu_{2D} < \frac{1}{3} \quad Less\ extensibility$$

$$For \quad \nu_{2D} > \frac{1}{3} \quad More\ extensibility$$

This is consistent with the limits at $\kappa = 0$, $\nu_{2D} = 1$ and at $\kappa = 1$, $\nu_{2D} = 0$. The latter case has already been discussed. The former case at $\nu_{2D} = 1$ has no resistance to bond-bending deformation (bond distortion), and represents behavior reminiscent of elastomers in their three-dimensional formalism, which allow very large deformations through flexibility of the high-molecular-weight polymer chains.

Further application of the moduli in (14.3) could involve evaluating the coefficient k_A directly from data or atomic-scale considerations, and then using that along with the other derived properties to examine the other mechanical properties. It should also be noted that the results in (14.2) and (14.3) are the initial, linear-range properties, whereas graphene is capable of extending into a considerable non-linear range of behavior; Lee *et al.* [14.1].

14.2 A Hypothetical Nanostructure

The model just derived and discussed for graphene suggests another closely related model for a different ideal, two-dimensional arrangement of atoms. Suppose each atom had six nearest-neighbors as in the arrangement shown in Fig. 14.4.

The basic repeating cell is triangular rather than hexagonal, but the hexagonal symmetry still applies at a larger scale. There is no known atomic arrangement of elemental atoms that gives this form, but it is an interesting, purely hypothetical nanoscale case to examine here. Each atom would require six bonding electrons (at least) in the outer shell of electrons.

Conducting exactly the same type of equilibrium analysis as that given above for graphene then gives the governing Poisson's ratio for this new arrangement of atoms as

$$\nu_{2D} = \frac{1-\kappa}{3+\kappa} \qquad \left(-1 \leq \nu_{2D} \leq \frac{1}{3}\right) \tag{14.10}$$

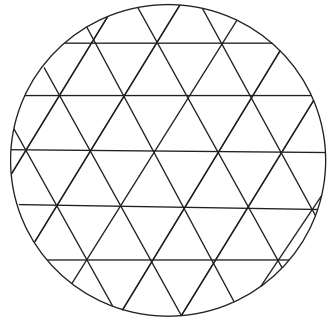


Fig. 14.4 *Triangular pattern of atoms.*

Table 14.1 Comparison of nanostructures

Nanostructure	Hexagonal	Triangular
ν_{2D}	$\dfrac{1-\kappa}{1+3\kappa}$	$\dfrac{1-\kappa}{3+\kappa}$
E_{2D}	$\dfrac{4\kappa}{\sqrt{3}\,(1+3\kappa)}k_A$	$2\sqrt{3}\left(\dfrac{1+\kappa}{3+\kappa}\right)k_A$
μ_{2D}	$\dfrac{\kappa}{\sqrt{3}\,(1+\kappa)}k_A$	$\dfrac{\sqrt{3}}{4}\,(1+\kappa)\,k_A$
K_{2D}	$\dfrac{k_A}{2\sqrt{3}}$	$\dfrac{\sqrt{3}}{2}k_A$

where again the limits on Poisson's ratio are found by allowing all possible non-negative values on k_A and k_B in (14.4).

Between the two nanostructures considered here, then, the complete range of Poisson's ratios from (14.2) and (14.10) is the usual and accepted range of

$$-1 \leq \nu_{2D} \leq 1$$

Now using the physical restriction (14.6) in (14.5), then (14.10) has the limits

$$0 \leq \nu_{2D} \leq \frac{1}{3} \tag{14.11}$$

for this new nanostructure.

The two nanostructures have the properties shown in Table 14.1 with the appropriate physical limits on κ given by (14.7). The significance of the results in the table, and thereby the uniquely special significance of graphene, will be discussed in the next section.

14.3 Comparison and Discussion

The first nanostructure in Table 14.1 is that of graphene, and the second is the purely hypothetical triangular nanostructure with no known atomic configuration.

From Table 14.1 it is seen that the bulk modulus K_{2D} is three times larger for the triangular nanostructure than it is for the hexagonal nanostructure at the same values of the bond-stretching stiffness coefficients. This is because of the greater number of theoretical atomic bonds in the former case. Now express the K_{2D}s as functions of the ν_{2D}s for the two nanostructures. Then express E_{2D} and μ_{2D} in terms of the ν_{2D}s for each nanostructure. These results for the like properties of the two nanostructures at the same Poisson's ratios are also found to differ by the same factor of 3 as do the K_{2D}s. Using Poisson's ratio in this manner provides a unification of the results. Poisson's ratio permits and provides a special scaling type of understanding of the behaviors and capabilities of these two nanostructures—one real and one hypothetical.

Despite the theoretical linkage described above, the comparison of the properties for the two nanostructures in Table 14.1 shows that they are fundamentally, irretrievably different. Atomic bonding in the case of carbon favors the first nanostructure, but the second nanostructure imposes barriers to atomic bonding. This makes the graphene nanostructure uniquely important, as it provides the only two-dimensional analog of what is possible with all the elements in their three-dimensional forms.

Now the question arises as to where graphene fits in explicitly with this range of conceivable behaviors that follow from this nanomechanics model. Poisson's ratio is difficult to measure directly, but there are clear indications of its probable value or narrow range of probable values for graphene. First of all, diamond is generally considered to have a Poisson's ratio of about $\nu = 0.20$. From Lee *et al.* [14.1] and Al-Jishi and Dresselhaus [14.2] there is reasonable evidence that for graphene the Poisson's ratio is somewhere in the range between 0.16 and 0.20. This then places graphene in the range having less extensibility than it would possess if it had a higher value of ν_{2D}, meaning lower values of κ and d/l. For the lower values of ν_{2D}, meaning higher values of κ, the bond-bending mechanism in this idealized elastic model suffers non-uniform stress conditions in the idealized beam-type members, and would be expected to fail at lower extensibilities.

So far, only the elastic moduli for these nanostructures have been examined. It is not at all clear how the explicit conditions of failure should be approached. For these cases of "perfect" materials with no defects or flaws, perhaps it is best to address the failure question through atomic-scale approaches such as density-functional theory. To the extent that

a macroscopic scale form for a failure criterion could be conceived, perhaps the best form for that of two-dimensional isotropic behavior would be the polynomial-invariants formulation for quasi-isotropic laminates. From (12.5) in Chapter 12 this failure criterion is

$$\left(\frac{1}{T}-\frac{1}{C}\right)(\sigma_{11}+\sigma_{22})+\frac{1}{TC}(\sigma_{11}+\sigma_{22})^2+\frac{1}{S^2}(\sigma_{12}^2-\sigma_{11}\sigma_{22})\leq 1 \quad (14.12)$$

where T, C, and S are the failure stresses in uniaxial tension and compression, and in shear. As before, the stresses must have units for force per unit length. Failure matters such as this are completely unexplored. It is generally understood that strength at the nanoscale is not the same critical issue that it usually is at the macroscale.

It is an intriguing question to ask what clues, if any, these two-dimensional ideal materials results may produce for the three-dimensional engineering materials of common use. To take a small step in that direction, examine the Poisson's ratios for those elements in the Periodic Table that take the form of solids. The maxima and minima are generally acknowledged to be:

Maximum ν	Gold Thallium Lead	0.44
Minimum ν	Beryllium	0.03

In these limit cases the maximum ν certainly conforms to extreme ductile behavior, while the minimum ν material is very brittle. This is consistent with the extensibility results found from graphene and its companion hypothetical nanostructure.

The case of beryllium provides a fascinating example. Brittle though it is, it has other spectacular properties at a very low density. If this material were not so virulently toxic and if it were in reasonable supply, it could well be the material of choice for an extremely broad range of applications. The present two-dimensional results show that the exceptionally low value of Poisson's ratio for beryllium probably involves much more of its resistance to deformation as resulting from bond-bending (distortion) than do other materials.

At the other end of the scale, of course, dislocations are responsible for ductile behavior, but only correlations with Poisson's ratio are being

considered in the present limited context. Thus at one extreme, gold has a very large Poisson's ratio $\nu = 0.44$, implying minimal resistive force due to bond-bending, while the other extreme, beryllium at $\nu = 0.03$, has the bond-bending resistance about as great as the bond-stretching resistance. In the latter case there results a very high level of stiffness, but at the expense of a very brittle failure behavior. Midway between these two extremes is carbon. For diamond with $\nu = 0.20$ and graphene with $\nu_{2D} = 0.20$ approximately, then from the stiffness coefficients (14.1) and with the corresponding $\kappa = 1/2$, the bond-bending resistance is half that of the bond-stretching resistance. The strength of the bond-bending resistance relative to the bond-stretching resistance is one of the fundamental controlling factors that profoundly influences materials behavior. Perhaps it is the single most important effect of all.

This probe into nanoscale properties provides an important example of the many opportunities that exist at that scale. Although nanoscale failure has little relationship to the usual mechanisms of macroscale failure, the nanoscale is still the logical place to start to understand all of the basic effects. None of the results found here could have been rationalized at the macroscale. In particular, the explicit relationship of the graphene physical properties and the bond-bending/bond-stretching effects only come into meaningful focus at the nanoscale. This recognition will be pursued further in the following section.

14.4 Are the Elements Ductile or Brittle?

The collection of the elements comprise the fundamental set of constituents, not only for all of nature but for the entire world of creative synthesis. The properties inherent in the individual elements are completely determinative in all of their fascinating combinations and applications. Although there is a broad array of relevant properties for the elements, two of the most important properties are stiffness and strength. When strength is brought into the picture there is always an adjoining consideration and qualification. Is the failure behavior ductile or brittle? This presentation is focused upon seeking a rational method for determining the relative ductility of the individual elements in the Periodic Table—that is, for any of those that form solids at ambient conditions. The term "relative ductility" means relative to the other elements, since there is some ambiguity about defining ductility in an absolute sense. Not surprisingly, it

will be found that the element gold provides the "gold standard" for this relative measure of ductility.

A more basic question than that of relative ductility concerns why some elements are ductile while some others are brittle. This can be answered only at the atomic or nanoscale, and it provides the guiding direction to be followed here. What characteristic(s) at the nanoscale determines whether a given element is ductile or brittle, and how does this lead to a method for quantitatively assessing all of the solids forming elements on this important issue?

The point of departure is that for the properties of graphene—the planar form that carbon atoms can take in forming a layer only a single atom in thickness. All bonding occurs with, and only with, the other atoms in the same plane. In this case of carbon, the atoms bond into a hexagonal pattern with the atomic centers at the nodes of the pattern. This case of graphene was studied in the previous section, and the method of analysis was as follows. The hexagonal pattern of carbon atoms was represented by a hexagonal pattern of two-dimensional elastic members possessing axial and bending stiffnesses. The two associated stiffness coefficients are representative of the bond-stretching and bond-bending capability between atoms.

With this mathematical formulation it was possible to derive the resultant in-plane elastic properties for the graphene arrangement of carbon atoms. One of the results was the consequent Poisson's ratio (14.2) repeated here for convenience and usage as

$$\nu_{2D} = \frac{1 - \kappa_{2D}}{1 + 3\kappa_{2D}} \tag{14.13}$$

where

$$\kappa_{2D} = \frac{\kappa_B}{\kappa_A} \tag{14.14}$$

and where k_A and k_B are the axial and bending stiffness coefficients for the hexagonal arrangement of elastic members. The elastic member, of unit thickness, had a width d and a length l between atomic centers. Then the non-dimensional nanoscale variable κ_{2D} was shown to be given by

$$\kappa_{2D} = \left(\frac{d}{l}\right)^2 \tag{14.15}$$

A particular value of κ_{2D} represents the specific identity of graphene. Further details are given in Section 14.1. These results for graphene are sufficient to proceed to the related but more general results.

Interest now is turned to the case of those elements in the Periodic Table that form solids at ambient conditions. In contrast to the two-dimensional state of graphene, the elements form fully three-dimensional states of matter (materials). Another contrast with graphene is that there is no specific crystalline form that produces isotropic conditions in the way that the hexagonal pattern in graphene produces two-dimensional isotropy. For the elements, the polycrystalline grains of various symmetries orient randomly to give a state of three-dimensional isotropy. Because of these complications it is not possible to derive directly the elastic properties for the elements from a realistic nanoscale architecture—at least not by the same method as that used for graphene. Accordingly, a different, less rigorous, but still quite closely related method will be followed here for the elements.

It is postulated that an expression for the Poisson's ratio of the elements still has the same form as that given in (14.13) for the two-dimensional model of graphene, but the coefficients embedded in it would be expected to possibly be very different from the two-dimensional case. Thus take for the elements

$$\nu = \frac{1 - \alpha\hat{\kappa}}{\beta + \gamma\hat{\kappa}} \tag{14.16}$$

where

$$\hat{\kappa} = \frac{k_B}{k_A} \tag{14.17}$$

In (14.16) α, β, and γ are coefficients to be determined, and $\hat{\kappa}$ in (14.17) is the controlling nanoscale variable expressed in terms of the axial and bending stiffness coefficients for the equivalent elastic members connecting atomic centers and simulating the bond-stretching and bond-bending resistances, corresponding to (14.14) and (14.15) for graphene.

Referring to a pair of bonded atoms, the bond-stretching mechanism involves the relative co-axial displacement of one atom with respect to the other. The bond-bending mechanism involves the relative tangential displacement between the two atoms. It is the nanoscale variable $\hat{\kappa}$ involving

these two mechanisms that will differentiate the various elements. The form (14.16), analogous to (14.13) for graphene, is thought to be well motivated, since it will be shown to be calibrated easily by certain known behaviors.

Before proceeding further, it should be noted that the bonding between neighboring atoms is usually but not always completely formed by the sharing of electrons in the outer shell of orbitals. The situation with the stiffness coefficients k_A and k_B is much more complicated. These coefficients represent the bond-stretching and bond-bending resistances of the atom as a whole, and as such have a dependence upon all the details of all the electronic orbitals and their interactions with each other in deformed configurations. Nevertheless, the aggregated effects embedded in k_A and k_B will provide valuable and sufficient information to proceed further.

Next, standard conditions will be used to evaluate α, β, and γ in (14.16). The following limits or anchor points must be satisfied

$$\nu = \frac{1}{2} \; at \; \hat{\kappa} = 0$$

and (14.18)

$$\nu = -1 \; at \; \hat{\kappa} = \infty$$

The first condition in (14.18) is that of no resistance to bond-bending, and the second case is that of infinite resistance to bond-bending. These conditions then reduce (14.16) to the form

$$\nu = \frac{1 - \alpha\hat{\kappa}}{2 + \alpha\hat{\kappa}} \tag{14.19}$$

One coefficient, α, remains to be determined. The form for the elastic members connecting atomic centers will be specified, as shown in Fig. 14.5.

The equivalent elastic member is of circular cylindrical form of radius r, diameter d, and length l. The stiffness coefficients for axial and bending stiffness are given by

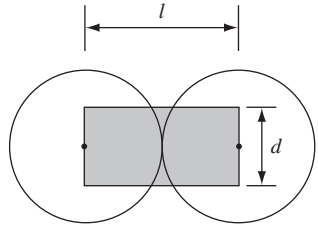


Fig. 14.5 *Equivalent elastic member between bonded atoms of the elements forming strong solids.*

$$k_A = \frac{A\tilde{E}}{l}$$
$$k_B = \frac{12\tilde{E}I}{l^3} \tag{14.20}$$

with

$$A = \pi r^2$$
$$I = \frac{\pi r^4}{4} \tag{14.21}$$

$\tilde{E}$ is the effective elastic constant for the equivalent elastic member. It should be noted that this entire procedure is mainly a mathematical construct—one that allows a consistent mathematical model to be developed. The concept of the equivalent elastic member is for this purpose only and not an actual, realizable entity.

Combining these specifications gives the nanoscale variable $\hat{\kappa}$ as

$$\hat{\kappa} = \frac{3}{4}\left(\frac{d}{l}\right)^2 \tag{14.22}$$

It is convenient to rescale the nanoscale variable $\hat{\kappa}$ as

$$\hat{\kappa} = \frac{3}{4}\kappa \tag{14.23}$$

So now

$$\kappa = \left(\frac{d}{l}\right)^2 \tag{14.24}$$

Finally, following exactly the same course as was found in the analysis of graphene, take

$$\nu = 0 \quad at \quad \kappa = 1, \quad \frac{d}{l} = 1 \tag{14.25}$$

This physical requirement states that the elastic member in Fig. 14.5 cannot extend beyond the outer shell of the electrons, and at this condition the related Poisson's ratio is zero. In the case of graphene this condition was derived, not assumed. As with graphene, negative values of ν would require that the equivalent elastic member extend beyond the outer shell of the electrons, $d/l > 1$. This would be physically unrealistic and unacceptable.

Using (14.22)–(14.25) gives $\alpha = 1$ in (14.19), leaving (14.19) as

$$\nu = \frac{1 - \kappa}{2 + \kappa} \tag{14.26}$$

This form is the end result of the derivation. It gives the macroscopic Poisson's ratio as a function of the nanoscale variable κ, where

$$\kappa = \frac{4}{3}\left(\frac{k_B}{k_A}\right) = \left(\frac{d}{l}\right)^2 \tag{14.27}$$

The limits on d/l are

$$0 \leq \frac{d}{l} \leq 1$$

and thereby on κ

$$0 \leq \kappa \leq 1$$

Inverting (14.26) gives

$$\kappa = \frac{1 - 2\nu}{1 + \nu} \tag{14.28}$$

This form will shortly be shown to be quite recognizable.

It is instructive to compare the three-dimensional result (14.26) with the corresponding two-dimensional result (14.13) derived for graphene. This comparison is shown in Fig. 14.6.

From Fig. 14.6 it is seen that there is something quite special about the value $\nu = \nu_{2D} = 1/5$ with both occurring at $\kappa = \kappa_{2D} = 1/2$, midway between the limits for the physically allowed levels for bond-bending and bond-stretching.

It is further interesting to compare this behavior with that for the stored energy in an elastic solid. The stored energy can be written as

$$U = \frac{1}{12\mu}\left[\left(\frac{1-2\nu}{1+\nu}\right)\sigma_{kk}^2 + 3s_{ij}\,s_{ij}\right] \tag{14.29}$$

where s_{ij} is the deviatoric stress and μ the shear modulus. It is seen that the coefficient of σ_{kk}^2 in (14.29) requires the value $\nu = 1/5$ when it is exactly half way between its allowed limits of 0 and 1. Thus the apportionment of energy at the macroscopic scale between dilatational and distortional types is completely in tune with the behavior of the nanoscale apportionment between bond-stretching and bond-bending. Compare the above form for U with the nanoscale result (14.28).

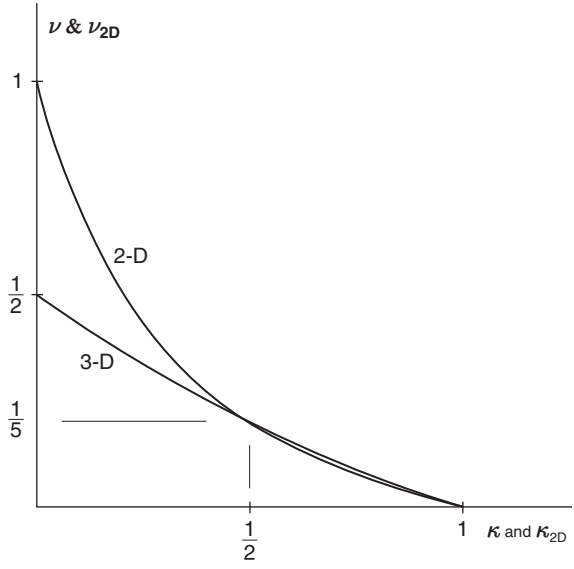


Fig. 14.6 *Comparison of two-dimensional and three-dimensional behaviors.*

The physical interpretation of the result (14.26) is as follows. The nanoscale variable κ ranges between 0 and 1. At $\kappa = 0$ there is no resistance from the bond-bending mechanism, and from (14.26) the corresponding Poisson's ratio is $\nu = 1/2$, allowing only shear deformation to occur. For any particular element, as the deformation increases, ultimately it causes yielding and/or failure. Near the $\kappa = 0$ limit the deformation can transition into shear localization or dislocation flow, or it is simply so great that ultimate failure becomes subsidiary to the excessive shear deformation. This is the case of dominant ductility.

At $\kappa = 1$, from (14.26) it is found that $\nu = 0$ and the behavior changes fundamentally from that at $\kappa = 0$. First note the form of the stress-strain relations at $\nu = 0$. In terms of principal stresses,

$$\begin{aligned} \sigma_1 &= E\varepsilon_1 \\ \sigma_2 &= E\varepsilon_2 \\ \sigma_3 &= E\varepsilon_3 \end{aligned} \tag{14.30}$$

There is no coupling whatsoever between these three relations. The failure behavior is consequently that of the maximum principal stress not exceeding a limiting value. Free-standing criteria of this type are generally taken to be reflective of brittle fracture behavior. This behavior at $\kappa = 1$ has the bond-bending resistance as about the same size as the bond-stretching resistance. There is no preferred mode of deformation rather than the necessary shear deformation that occurs at and near $\kappa = 0$. This is why the mode of failure changes so drastically between $\kappa = 0$ and $\kappa = 1$.

The failure type expressed as the degree of ductility is taken to vary continuously with the variation of κ, going from extremely ductile to extremely brittle over the range from 0 to 1. It follows from the result (14.26) that the rank order list of the elements in terms of ascending κs is the same as the rank order list of the elements in terms of descending νs. Finally, using the Poisson's ratios reported from the testing of the elements permits the construction of the rank order list of the ductilities for the elements. This list is shown in Table 14.2 for the more commonly known and used elements that form solids.

The table gives the relative ranking of ductility levels for the elements shown by the method developed here, with all cases referring back to the element gold. This method relates the observable macroscopic property ν to the controlling nanoscale variable κ. As such this rank ordering goes

Table 14.2 Properties for the elements

Element	ν	Nanoscale κ	Ductility $(1-\kappa)^2$
Absolute limit, perfect ductility	1/2	0	1
Gold	0.44	0.08	0.84
Lead	0.44	0.08	0.84
Niobium	0.40	0.14	0.73
Palladium	0.39	0.16	0.71
Platinum	0.38	0.17	0.68
Silver	0.37	0.19	0.66
Vanadium	0.37	0.19	0.66
Tin	0.36	0.21	0.63
Aluminum	0.35	0.22	0.60
Copper	0.34	0.24	0.58
Tantalum	0.34	0.24	0.58
Titanium	0.32	0.27	0.53
Cobalt	0.31	0.29	0.50
Nickel	0.31	0.29	0.50
Magnesium	0.29	0.33	0.45
Iron	0.29	0.33	0.45
Tungsten	0.28	0.34	0.43
Zinc	0.25	0.40	0.36
Manganese	0.23	0.44	0.31
Uranium	0.23	0.44	0.31
Silicon	0.22	0.46	0.29
Plutonium	0.21	0.48	0.27
Chromium	0.21	0.48	0.27
Carbon (diamond)	0.20	0.50	0.25
Limit for most elements/materials	1/5	1/2	1/4
Beryllium	0.032	0.91	0.01
Absolute limit, total brittleness	0	1	0

by that of the Poisson's ratio values. Most of the values for ν are taken from "Poisson's Ratio of the Elements" [14.3], as compiled for the online resource *Mathematica* from a variety of sources. Only three of the elements shown—carbon, silicon, and manganese—were not available from this reference. The values for the polycrystalline form of these elements were obtained from other sources that appear to be reliable, although there necessarily is some uncertainty.

14.5 A Ductility Scale for the Elements

In basic terms the table of ductilities, Table 14.2, is obtained from the nanoscale variable κ which characterizes the relative size of the bond-bending and bond-stretching effects. These effects are shown to relate to ductile and brittle failure behaviors. Finally, the relationship between κ and Poisson's ratio allows the actual assembly of the table. A further consequence of the bond-bending/bond-stretching behavior κ is that in general the more ductile elements with small κ have lower values of overall macroscopic stiffness, whereas the more brittle elements with large κ's have relatively large stiffnesses. These nanoscale effects provide the explanation of why there is a "trade-off" between high stiffness and high ductility for the elements. Presumably the same can be said of the strength levels.

It must be cautioned that determining Poisson's ratio accurately is a difficult proposition, and there could be considerable uncertainty in some of these reported values for ν. Also, it must be remembered that these are reported values for the elements, not for commercial materials, often of the same common name but with much more complex formulations. Nevertheless, the relative ductility's shown in Table 14.2 are in reasonable and general accord with common perceptions and observation of ductile-versus-brittle behaviors.

The third entry in Table 14.2 is that of $(1 - \kappa)^2$. The nanoscale variable κ goes from 0 to 1, so $(1 - \kappa)$ goes from 1 to 0. Expressing that in a quadratic form, somewhat like energy and some failure forms, then gives a qualified guide to the level of ductility for all the elements in the table, going from perfect ductility at 1 to no ductility, total brittleness at 0. According to this measure of ductility the division between or transition from the ductile elements to the brittle elements occurs about at the elements of cobalt and nickel with $(1 - \kappa)^2 = 0.5$, midway between the theoretical limits.

This provisional ductility scale at least gives some sense of the actual ductility levels rather than just the relative rankings. From Table 14.2 it is clear that only gold and lead are in the extremely ductile class, while only carbon (diamond) is probably in the very brittle class and beryllium is extremely brittle. One could proceed one step further and transform the ductility versus κ in the third column of Table 14.2 into an S-shaped curve that would have large sensitivity to increments in κ in the middle range, but much less sensitivity approaching the upper and lower limits of κ. Although this could be a more suggestive form than that in Table 14.2, it would be no more basic.

As already mentioned, and as evident in Table 14.2, it appears that the value of the nanoscale variable $\kappa = 1/2$ with the corresponding macroscale variable $\nu = 1/5$ has a very special significance, both for the elements and possibly for all isotropic materials. Beryllium predominates in the table as a special case—very probably due to its extremely low atomic number. Even so, it is completely consistent with this general ductile-versus-brittle classification.

The twenty-five elements in Table 14.2 are located all over in the Periodic Table, although sixteen of them are from the transition-metals category. The ductility rank ordering in Table 14.2 does have some correlation with the numbers of shells containing electrons (electronic shells) in the atomic configurations. First, it is observed that the two most ductile elements—gold and lead—each have six electronic shells, while the two most brittle elements—beryllium and carbon—each have but two shells. The two elements at the transition between the ductile and the brittle elements in Table 14.2—cobalt and nickel—each have four electronic shells.

The actinides, including uranium and plutonium, are in a separate category by themselves. If two further elements tantalum, and tungsten are temporarily eliminated from consideration, then the remaining twenty one elements in Table 14.2 show a quite good correlation between increasing ductility and increasing numbers of electronic shells. This apparent correlation means that increasing the number of electronic shells decreases the nanoscale variable κ that specifies the ratio of the bond-bending resistance to the bond-stretching resistance. It is not unexpected that these physical characteristics at the atomic and nano scales could be strongly related.

No doubt many other factors also are necessarily involved to account for each and every individual element. It remains an open question whether any further relationships can be established linking the present

rank ordering of ductility's with any of the electronic properties of the orbitals for the elements. There does not appear to be any simple and direct relationship to the valence electrons.

It must be emphasized that this ordering of the ductilities is only for the elements, not for all materials. For more general materials the use of Poisson's ratio as an indicator of relative ductility levels must be used with caution. Although the procedure may be somewhat useful in general, there would be many important exceptions where the ductilities depend upon many factors other than just the relationship of the bond-bending capability to the bond-stretching capability for a single constituent in a compound containing many different elements. Flaws and defects would be of decisive importance in general materials.

When one investigates the incredibly broad category of all homogeneous materials, this use of Poisson's ratio to estimate ductility should be replaced by a more finely tuned method for studying ductile and brittle failure. For one thing, acknowledgment must be given to the importance of the stress state under consideration. The present approach for the elements probably applies only to the state of uniaxial tension. In Chapter 8, all matters related to ductile and brittle failure are taken up and accessed through the materials type based upon its T/C ratio, where T and C are the strengths in uniaxial tension and compression, and also where all stress states are considered.

Despite the limitations just discussed as qualifications of this method of characterizing ductile-versus-brittle behavior, it still provides a remarkable view of the power of nanoscale analysis. In particular, this method has succeeded in relating a non-destructive macroscopic property—Poisson's ratio—to some very important aspects of failure behavior. In so doing it has been found that Poisson's ratio is not just a minor adjunct to Young's modulus, E, as is often gratuitously implied.

The properties E and ν are equal partners in specifying the controlling properties for all isotropic materials. The dimensional property E mirrors the load-bearing capability of the material, while non-dimensional ν has subtle and vital implications concerning the underlying nanostructure of the material, especially as regards matters of ductile and brittle failure. To express this another way: for the elements, modulus E represents the aggregated, combined effect of the resistances from atom-to-atom bond-stretching and bond-bending, while ν represents and reflects the differences between the bond-bending and bond-stretching mechanisms.

Problem Areas for Study

1. The present failure theory for isotropy is for macroscopic field theory applications. It is calibrated by uniaxial T and C. Can these two macroscopic strength properties be expressed in terms of nano (or micro) scale variables to generate a nano (or micro) scale theory of macroscopic strength?
2. Can nanoscale indentation be used to critically assess macroscopic failure criteria?
3. In the graphene analysis in Section 14.1 the properties are interpreted to some degree in terms of tensile failure behavior. Is there any useful interpretation for compressive failure at the nanoscale, particularly for graphene?
4. Why is the hypothetical nanostructure in Section 14.2 only a concept and not a real entity formed by some element or elements? What atomic/orbital properties prevent realization of this hypothetical nanostructure? If this nanostructure were to exist, would it be inferior or superior to graphene?
5. Section 14.5 deduces a ductility scale for those elements that form strong solids. It might be helpful to revise the scale such that it more nearly reflects the actual ductility of a particular element relative to the entire group of them. For example, a form that has the greatest rate of change at the center of the list, and the least rate of change near the upper and lower limits, would probably be realistic. The rank ordering of the respective ductilities must remain unchanged. Can this change of scale be achieved, and if so, how?

References

[14.1] Lee, C., Wei, X., Kysar, J. W., and Hone, J. (2008). "Measurement of the Elastic Properties and Intrinsic Strength of Monolayer Graphene", *Science* **321**, 385–8.

[14.2] Al-Jishi, R., and Dresselhaus, G. (1982). "Lattice–Dynamical Model for Graphite", *Physical Review* **B26**, 4514–22.

[14.3] "Poisson's Ratio of the Elements" (2012). http://periodictable.com/Properties/A/PoissonRatio.html

15

Damage, Cumulative Damage, Creep and Fatigue Failure

Damage in materials is pervasive and pernicious. Damage occurs both in the manufacturing/synthesis stage and in the lifetime structural usage stage. Normally materials with manufacturing defects are treated as virgin materials that necessarily have a performance level which is less than that of the ideal, perfect product. The second type of damage—that which occurs in applications—is the source of most of the difficulties, and is of major interest here as it relates to failure. After an introduction to the general subject of damage, the main topic of cumulative damage will be taken up. Cumulative damage is concerned with the methods by which a sequence of damage events are taken into account and combined, ultimately resulting in the failure of the material. Specialized damage in fiber composite laminates was treated in Chapter 14. Now, damage in general materials will be treated, with most applications being intended for isotropic materials.

15.1 Damage

For initial purposes it is helpful to consider three distinctly different types of damage in materials. First, there is the type of damage that is readily apparent before major yielding/failure. Second, there can be apparent damage that occurs after the first major yielding, but during the continued loading phase. Third, there can be the type of barely observable, extremely small-scale damage that occurs during loading but with no apparent consequences until abrupt final failure. These three cases, in different order, are exemplified by (i) damage after yielding in ductile metals, (ii) transverse matrix cracking before final failure in fiber composites, and (iii) subscale damage preceding fatigue failure in most materials.

The three cases stated above will be expanded in the following discussion. Of course, these cases are somewhat idealized and actually can

sometimes show features composed of a complex combination of these. We begin with the case of ductile metals, which can fail in the mode called "ductile fracture." The two words "ductile" and "fracture" in combination almost seem to pose a contradiction. Nevertheless, they do describe a reality that is related to and is usually caused by damage.

If ductile fracture does not initially occur, then the material simply continues to undergo plastic flow in a ductile manner until something else intercedes. After the point of major yielding and during the strain hardening stage of loading, voids in the material can form and grow. A well-known model for the effect of spherical voids upon ideal yielding is that due to Gurson, as given by

$$\frac{3}{2}s_{ij}\,s_{ij} + 2v\,T^2\left[\cosh\left(\frac{\sqrt{3}}{2\,T}\sigma_{ii}\right) - 1\right] = (1-v)^2\,T^2$$

where v is the volume fraction of voids, and T is the yield stress in uniaxial tension (and compression). For $v = 0$ the form is that of the Mises criterion where s_{ij} is the deviatoric stress tensor. The presence of the voids as well as a possible mean normal stress effect, σ_{ii}, acting with the voids, diminishes the yield stress from what it would be for a Mises criterion. Much more detailed forms have been devised using Gurson's form as the starting point. These are as given by Tvergaard, Needleman, and others, and reasonably general approaches by Nahshon and Hutchinson [15.1] and by Xue [15.2] contain references to the earlier works. In all these works criteria are built into the procedure that specify the nucleation of the voids, the rate of growth of them, the interaction of them, and the possible coalescence of them to ultimately cascade into failure. These treatments also bring in the third invariant, not just the first two invariants as in the Gurson form. This circumstance is specified through a dependence upon the Lode angle in principal stress space that implies the existence and utilization of non-convex yield surfaces. All the effects just described, possibly leading to ductile fracture in ductile metals, involve inherently inhomogeneous effects, and they dramatically illustrate the difficulties and complications that can follow from this type of void growth damage.

In contrast to the above type of analysis, in the homogeneous material idealizations of this book the point of major yielding is interpreted as that

of yield/failure. This approach, then, does not deal with the complication of whether or not the plastic flow is interrupted by fracture. The major load-bearing function has already been sacrificed at the point of major yield, which is thus taken as that of effective failure.

In further contrast to the above type of damage, fiber composite materials can undergo very extensive transverse cracking between fibers without showing much effect on the apparent stress–strain behavior. This primarily occurs in polymeric-matrix carbon-fiber systems, which are the most important type of high-performance composite. However, the apparently benign nature of this type of damage can be deceiving. Transverse cracking within a lamia—which lamina is itself within a laminate—can provide the nucleation sites and energetics for delamination of the laminate. Delamination is the major weakness element in these materials. The sequencing effect of these sources of degradation in these types of materials are often called "progressive damage."

This book has two chapters concerned with the failure of fiber composite materials. The first is at the aligned fiber, lamina level, and represents the rather idealized failure type without damage at that level. Then failure at the laminate level is treated, involving the interactive effects of the ideal failure at the lamina level and the effect of progressive damage at the laminate level, ultimately resulting in the complete failure of the laminate.

Yet another type(s) of damage must be involved under the cyclic loading conditions of fatigue. In this condition the damage is not readily discernable—at least not in a macroscopic sense. Yet the damage relentlessly grows and spreads until enough damage has accumulated to lead to sudden and often almost instantaneous failure. In materials with grains this could be localized failure in the grain boundaries, while in fiber composites it could be debonding between the fiber and matrix phases. There is an unlimited number of different possible damage mechanisms for the different materials types.

At first exposure it would seem almost impossible to theoretically characterize the growing state of fatigue damage along its transit toward failure. Nevertheless, out of necessity several approaches have been developed. Most of these are phenomenological in the sense that they are simply postulated forms that can be tested in different ways, while a few approaches are physical in the sense that they represent a state of preexisting cracks subject to kinetic growth control. It is this latter approach

that will be mainly expanded in this chapter, although representative cases of phenomenological forms will also be examined.

This fatigue type of damage modeling is given under the heading of cumulative damage. It is found that a formalism can be taken that applies to either fatigue or creep rupture through suitable notational changes. The most basic problem associated with this area of study is to predict the lifetimes under variable amplitude loading conditions based upon knowledge of the database of lifetimes under constant amplitude loading conditions. Other than the variation in the amplitude, all other characteristics such as frequency are held constant. In this manner the lifetime can then be predicted for any specified history of variable load amplitudes.

These three major types of damage at least show the complexity of the microscale events leading to failure. Failure at the atomic-bond level or at the nanoscale or the microcscale do not necessarily imply or commit to failure at the macroscale. Of course, failure at the macroscale does imply all of the above, but failure at the macroscale itself has special meaning and restrictions. Nevertheless, the quantification of damage states can sometimes provide the link between all these irreversible and deleterious subscale events and the final, ultimate failure of the material.

15.2 Cumulative Damage

Of the various headings under mechanics of materials, failure is usually considered to be the most difficult. Then, proceeding further to the subfields under failure, fatigue is probably considered to be its most difficult one. The concept of damage comes into play with fatigue, but deducing useful and reasonably general damage forms has been a taxing and uncertain exercise. The organized attempts to quantify growing states of damage usually falls under the heading of cumulative damage, with the understanding that the damage leads to and terminates with materials failure.

Quantifying damage relative to failure allows the optimization of stress-loading programs and allows the means for including overload-type damage events. Although this book is not intended as a literature survey, this section will be used to examine several different approaches for cumulative damage. The possible applications are to isotropic or anisotropic

materials (including composites) with the proviso that in either case the conditions of proportional loading apply.

Creep rupture also comes under the heading of cumulative damage. It is the time-dependent damage growth leading to failure in polymers and in metals at high temperature. As developed here, fatigue conditions and creep rupture conditions admit the same general formulation. Both will be treated, and with suitable notational changes either form can be found from the other.

There are considered to be two main approaches for cumulative damage. One is that of the direct postulation of lifetime damage forms such as Miner's rule, and the other is that of residual strength. Residual strength is the reduced (instantaneous) static strength that the material can still deliver after being subjected to loads causing damage. Of course, both are relevant to the general problem, and a well-founded cumulative damage theory must contain both descriptions in a compatible form. The overall objective of both and all approaches is to secure a life prediction methodology.

In the area of the fatigue of metals, a common approach is to assume a pre-existing crack that grows according to a power-law form with regard to the stress level. This is often called the "Paris law." When the crack reaches a certain prespecified size, the service life is considered to be completed. While this is certainly a useful approach, it cannot be considered to be a life prediction methodology based upon the approach to failure. Accordingly, it will not be followed here.

The general topics of fatigue or of damage are covered in many self-contained, inclusive books such as Suresh [15.3] for fatigue and Krajcinovic [15.4] for damage. As part of a general damage approach, constitutive relations are often taken for damage involving scalar- or tensor-valued damage variables, and a whole framework of behavior is built up. In a considerably different direction to be followed here, cumulative damage leading to failure is more akin to failure criteria, but with the capability to characterize damage as the prelude to failure. The first credible damage form was that of Miner's rule [15.5]. Broutman and Sahu [15.6] and Hashin and Rotem [15.7] much later produced other damage forms that with the passage of time have gained credibility. Reifsnider [15.8] initiated a different damage formalism that has been further developed by him and by others in various papers. Other also notable

efforts have been presented by Adam *et al.* [15.9] and by many other workers. Christensen [15.10] developed a kinetic crack growth approach. Post, Case, and Lesko [15.11] have given a survey of many different cumulative damage models and applied them to situations of spectrum-type loadings for composites.

Due to the complexity of the topic, many approaches and models contain adjustable parameters—sometimes many parameters. In the coverage undertaken here, only models without any adjustable parameters will be considered. That is to say, the damage formulations to be considered will be based solely upon the properties contained within a database of constant amplitude fatigue or creep testing and the static strength. Then the damage forms will be used to predict life under variable-amplitude programs of load application. This circumstance is analogous to viscoelastic behavior where mechanical properties of creep or relaxation function type are inserted into convolution integrals to predict general behavior. Materials with memory could encompass viscoelastic effects at one extreme, and damage memory effects at the other extreme. The four damage models to be examined here are those of Miner's rule, Hashin–Rotem, Broutman–Sahu, and then the kinetic crack growth method developed by Christensen. The Broutman–Sahu form is representative of a large class of models, as will be explained later.

In terms of evaluating damage models, there does not appear to be any definitive set or sets of experimental data. The problem is the lack of repeatability, and in a more general sense the great variability in the data. This places even more emphasis on the need for a careful theoretical evaluation, looking for physical consistency in the predictions and conversely examining aspects of inconsistency in important special cases.

Three main problems will be used to evaluate the four models. These include examples of prescribed major damage followed by the predicted life at a lower, safer stress level. Another problem is that of the ability to predict residual instantaneous strength after a long time load application that nevertheless is shorter than that time which would cause fatigue or creep failure. Third is the problem of residual life. In this situation a load is applied right up to the time of, but just an instant before, fatigue or creep failure. At that time the load is removed and replaced by one of a lower, less stringent level. The additional life that results at the lower stress level is designated as the residual life. This detailed collection of predictions will present a comparative theoretical evaluation of the four models.

15.3 Four Models

The general concept of damage that would or could lead to failure was rather vague and nebulous until Miner gave it a specific form and it became recognized. With background from Palmgren, Miner [15.5] postulated the damage form for fatigue as

$$\sum_i \frac{n_i}{N_i} = 1 \tag{15.1}$$

where n_i is the number of applied cycles at nominal stress σ_i, and N_i is the limit number of cycles to failure at the same stress and for the same cycle type. Thus each value n_i/N_i is viewed as a quantum of damage, the sum of which specifies failure. As with all cumulative damage forms, when the left-hand side of (15.1) is less than 1, it still quantifies the damage level but does not imply failure. The spectrum of values of N versus constant stress, σ, is as shown in Fig. 15.1, where for constant stress $n = N$. Relation (15.1) then allows the life prediction for a combination of different load levels. All of the fatigue conditions considered here will be taken to be of the same frequency and cycle type.

Relation (15.1) is a completely empirical form, but was a reasonable conjecture at the time. It is usually called Miner's rule, but sometimes goes by the name Palmgren–Miner rule or law, and is also termed "linear cumulative damage".

The first evidence of the possible inadequacy of Miner's rule is that it predicts the independence of the order of application of the loads leading

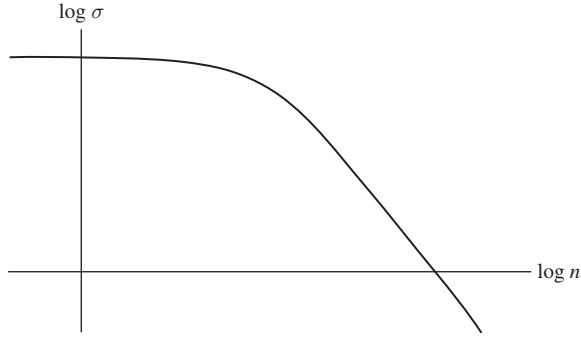


Fig. 15.1 *Fatigue life at constant stress.*

to failure, so long as the duration of each sequence of cycles is preserved. Despite the shortcomings which will be shown in the following evaluation, Miner's rule has always been by far the most widely used cumulative damage form.

Writing Miner's rule in its two-step form gives

$$\frac{n_1}{N_1} + \frac{n_2}{N_2} = 1 \tag{15.2}$$

Hashin and Rotem [15.7] recognized that there could be a problem with the calibration of the first term relative to the second. Specifically, for a given proportion of the life being expended in step 1 at stress level σ_1, this proportion of the expended life would be expected to be different relative to the following stress level of step 2. With a set of conditions and assumptions they modified (15.2) to the form

$$\left(\frac{n_1}{N_1}\right)^{\frac{\log N_2}{\log N_1}} + \frac{n_2}{N_2} = 1 \tag{15.3}$$

For more than two steps, the form (15.3) can be applied iteratively, and can also be modified to accommodate a continuous variation of n/N, though the result is quite complex and is not needed here.

In a completely different approach the problem of residual strength can be approached. Residual strength is the static strength that would exist if a fatigue test were interrupted and tested to failure in the usual static strength manner. One of the simplest possible residual strength forms was presented by Broutman and Sahu [15.6]. It states the residual strength σ_R as

$$\sigma_R = \sigma_s - \sum_i (\sigma_s - \sigma_i)\frac{n_i}{N_i} \tag{15.4}$$

where σ_s is the instantaneous tensile static strength as measured on the virgin material before any fatigue damage is induced. At $\sigma_i = \sigma_s$ (with σ_i being the maximum stress in the cycle) this residual strength form properly reduces to the static strength. The form (15.4) is emblematic of a much larger class of models whereby $(\sigma_s — \sigma_i)$ and (n_i/N_i) are each raised to some different power, thus introducing into the process adjustable parameters to be determined.

The life of the material under a prescribed stress program is implicit in (15.4) and is determined by the total number of cycles at which $\sigma_R = \sigma_i$, the nominal (maximum) applied stress and the residual strength become identical. Before this number of cycles the residual strength is greater than the applied maximum stress but less than the static strength.

At this point it is convenient to recast these fatigue forms into analogous creep rupture forms. Creep rupture occurs in polymers and in metals at high temperature. Let $t_c(\sigma)$ be the time to failure under constant stress σ. Miner's rule (15.1) for creep conditions then takes the form

$$\int_0^{\tilde{t}} \frac{d\tau}{\tilde{t}_c(\tilde{\sigma})} = 1 \tag{15.5}$$

where stress is non-dimensionalized by the static strength

$$\tilde{\sigma} = \frac{\sigma}{\sigma_s}$$

and time is non-dimensionalized by a time constant, $t_{0,}$ that calibrates the time scale,

$$\tilde{t} = \frac{t}{t_0}$$

The basic property $\tilde{t}_c(\tilde{\sigma})$ gives the failure times versus the constant stress levels, as shown in Fig. 15.1, but for creep rupture rather than fatigue.

The Hashin–Rotem fatigue form (15.3) then for a two-step creep condition becomes

$$\left(\frac{\tilde{t}_1}{\tilde{t}_{c1}}\right)^{\frac{\log\ \tilde{t}_{c2}}{\log\ \tilde{t}_{c1}}} + \frac{\left(\tilde{t} - \tilde{t}_1\right)}{\tilde{t}_{c2}} = 1 \tag{15.6}$$

where

$$\tilde{t}_{c1} = \tilde{t}_c(\tilde{\sigma}_1)$$
$$\tilde{t}_{c2} = \tilde{t}_c(\tilde{\sigma}_2)$$

The Broutman–Sahu residual strength fatigue form (15.4) becomes the creep condition residual strength as

$$\tilde{\sigma}_R = 1 - \int_0^{\tilde{t}} \frac{(1-\tilde{\sigma})\,d\tau}{\tilde{t}_c(\tilde{\sigma})} \tag{15.7}$$

The corresponding lifetime form for Broutman–Sahu is found by taking $\tilde{\sigma}_R(t) = \tilde{\sigma}(t)$ in (15.7) to obtain

$$\frac{1}{1-\tilde{\sigma}(t)} \int_0^{\tilde{t}} \frac{(1-\tilde{\sigma})\,d\tau}{\tilde{t}_c(\tilde{\sigma})} = 1 \tag{15.8}$$

Compare (15.8) with the Miner's rule form (15.5). Considerable differences must be expected.

Now the fourth model will be introduced. In a program to develop a physically based flaw growth model, Christensen [15.10] has obtained a completely different formalism. Since a flaw growth method cannot simply be postulated, its derivation will be outlined here in a brief manner.

Take an existing microscale crack and specify its rate of growth according to a power law as

$$\dot{a} = \lambda(\sigma\sqrt{a})^r \tag{15.9}$$

where the crack size $a(t)$ and the stress $\sigma(t)$ are functions of time, r is the power-law exponent, and λ is a constant.

Integrate (15.9) to obtain

$$\left(\frac{a}{a_0}\right)^{1-\frac{r}{2}} - 1 = \lambda\left(1-\frac{r}{2}\right) a_0^{\frac{r}{2}-1} \int_0^t \sigma^r(\tau)\,d\tau \tag{15.10}$$

where a_0 is the initial crack size.

Following classical fracture mechanics as an initial approach, take the crack as growing to a size that becomes unstable when it reaches the same stress intensity factor as that which gives the static strength, σ_s, thus

$$\sigma(t)\sqrt{a(t)} = \sigma_s\sqrt{a_0} \tag{15.11}$$

Combining (15.10) and (15.11) then gives a life prediction form to be solved for the lifetime t under a prescribed stress history $\sigma(t)$. When this form is specialized to the constant stress case, the creep rupture time to failure is found to be given by

$$\tilde{t}_c = \frac{1}{\tilde{\sigma}^r} - \frac{1}{\tilde{\sigma}^2} \tag{15.12}$$

where non-dimensional stress and non-dimensional time will be used from this point on. Relation (15.12) has a behavior as shown in Fig. 15.2a on log–log scales. At long time there is a power-law behavior controlled by the exponent r, and at short time it approaches the static strength asymptote. That much is perfectly acceptable. However, the form (15.12) exhibits a rather sharp transition from one asymptote to the other. Most data show a more gradual transition. Thus a more general approach may be required, although the form (15.12) could remain as a special case.

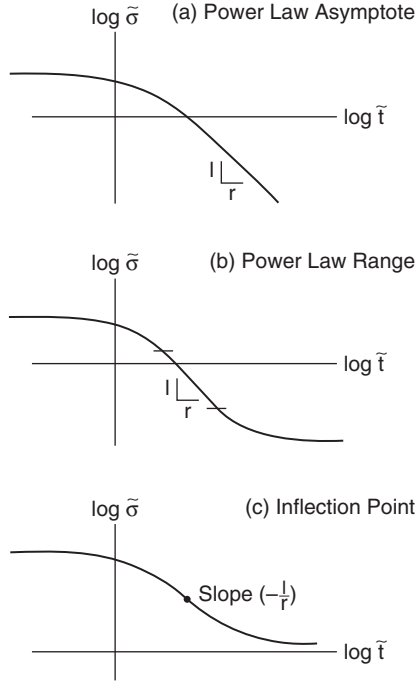


Fig. 15.2 *Determination of exponent r from creep rupture at constant stress.*

Going back to relation (15.11) that gives the criterion for unstable growth of the crack, it will now be generalized. Instead, replace (15.11) by the more general form

$$\tilde{a} = F(\tilde{\sigma}) \tag{15.13}$$

where

$$\tilde{a} = a(t)/a_0$$

and $F(\tilde{\sigma})$ is a function of stress yet to be determined. The reason for this generalization beyond the behavior of a single ideal crack relates to more complex matters such as the possible interaction between cracks, crack coalescence, and many other non-ideal types of damage and damage growth. The basic flaw growth relation (15.9) will be retained, but now $a(t)$ represents some more general measure of flaw size as it grows. The method will still be referred to as kinetic crack growth. The types of behavior to be represented by this more general approach are as shown in Fig. 15.2.

Combining (15.10) and (15.13) gives

$$f(\tilde{\sigma}(\tilde{t})) = \int_0^{\tilde{t}} \tilde{\sigma}^r(\tau) d\tau \tag{15.14}$$

where

$$f(\tilde{\sigma}) = 1 - F^{1-\frac{r}{2}}(\tilde{\sigma}) \tag{15.15}$$

and where the various properties combine to form the calibrating time constant.

For constant stress, relation (15.14) becomes

$$f(\tilde{\sigma}) = \tilde{\sigma}^r \tilde{t}_c(\tilde{\sigma}) \tag{15.16}$$

where $\tilde{t}_c(\tilde{\sigma})$ is the spectrum of creep rupture properties at stress levels $\tilde{\sigma}$, taken to be known from tests.

Finally, substituting (15.16) into (15.14) gives the flaw growth lifetime criterion as

$$\frac{1}{\tilde{\sigma}^r(\tilde{t})\tilde{t}_c(\tilde{\sigma})}\int_0^{\tilde{t}}\tilde{\sigma}^r(\tau)d\tau = 1 \tag{15.17}$$

For a given stress history $\tilde{\sigma}(\tau)$, relation (15.17) determines the lifetime, $\tilde{t}$, of the material. The exponent r is determined from the basic creep rupture forms as shown in Fig. 15.2. The creep rupture properties and its specific property r calibrate the theory behind the life-prediction form (15.17). It is not surprising that the property $\tilde{t}_c(\tilde{\sigma})$ in (15.17) is at current time, $\tilde{t}$, rather than being inside the integral as in the other models. This relates to the method whereby the flaw grows until it reaches a critical size at the then existing current stress. This present approach also admits a full statistical generalization that will be taken up in the next and final chapter.

The four basic forms under consideration here—Miner's rule, Hashin–Rotem, Broutman–Sahu, and the kinetic crack form—are respectively given by (15.5), (15.6), (15.8), and (15.17), It is important to observe that all of these are completely calibrated by and determined by only the basic experimentally determined creep rupture property $\tilde{t}_c(\tilde{\sigma})$ and the static strength. There are no additional parameters to be adjusted or fine-tuned. Apparently these four are the only forms that have been proposed or derived that do not involve additional parameters beyond the mechanical properties.

It is interesting to note that there is a special case in which two of these basic forms become identical. Specifically for a creep rupture property of the power-law type, as in

$$\tilde{t}_c = \frac{A}{\tilde{\sigma}^r} \tag{15.18}$$

then Miner's rule (15.5) and the kinetic crack form (15.17) reduce to the same form. This is shown directly by substituting (15.18) into each of these. Otherwise these two forms make completely different predictions—sometimes extremely different predictions. In this power-law special case, (15.18), the creep rupture conforms to the limiting (degenerate) case shown in Fig. 15.3.

The residual strength form corresponding to the kinetic crack form (15.17) is given by

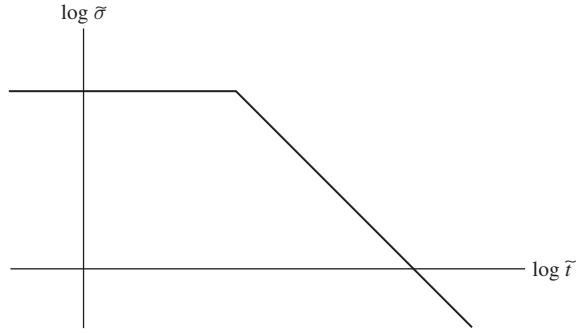


Fig. 15.3 *Creep-rupture power-law behavior.*

$$(\tilde{\sigma}_R)^r \tilde{t}_c(\tilde{\sigma}_R) = \int_0^{\tilde{t}} \tilde{\sigma}^r(\tau) d\tau \tag{15.19}$$

The residual strength form of the Broutman–Sahu type is given by (15.7). The other two models do not directly give a residual strength determination.

The fatigue life forms for the first two models are given by Miner's rule (15.1) and Hashin–Rotem (15.3). The Broutman–Sahu form for fatigue life is given from (15.4) by setting $\sigma_R = \sigma_i$. The fatigue life form from the kinetic crack derivation is given by

$$\frac{1}{(\sigma_K)^r N(\sigma_K)} \sum_{i=1}^{K} (\sigma_i)^r n_i = 1 \tag{15.20}$$

where exponent r is given by the basic fatigue property envelope at constant amplitude—the same as in the creep rupture curves of Fig. 15.2.

This completes the background and specification of the four basic forms under consideration. Now a comparative evaluation of these forms will be given using the creep failure notation.

15.4 Residual Strength

When a general program of stress history is interrupted and suddenly tested for static strength, Fig. 15.4, the residual strength, σ_R, is determined.

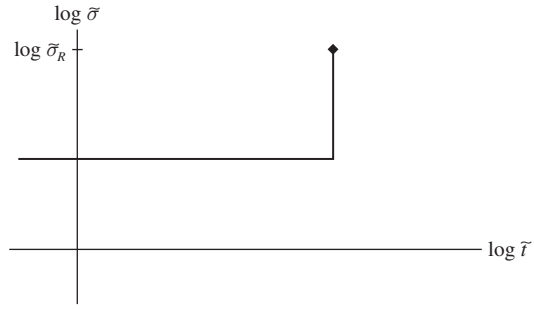


Fig. 15.4 *Residual strength.*

Under constant stress $\tilde{\sigma}$, before testing for $\tilde{\sigma}_R$ at time $\tilde{t}_1$ the Broutman–Sahu form (15.7) becomes

$$\tilde{\sigma}_R = 1 - (1 - \tilde{\sigma})\frac{\tilde{t}_1}{\tilde{t}_c(\tilde{\sigma})} \tag{15.21}$$

Under constant stress until time $\tilde{t}_1$, the kinetic crack form (15.19) becomes

$$(\tilde{\sigma}_R)^r \tilde{t}_c(\tilde{\sigma}_R) = \tilde{\sigma}^r t_1 \tag{15.22}$$

To determine $\tilde{\sigma}_R$ from (15.22) requires knowledge of the range of values for $\tilde{t}_c(\tilde{\sigma}_R)$. In contrast, (15.21) only requires knowledge of the value of $\tilde{t}_c$ at one value of stress. This difference has important implications, as will be seen next.

Consider the two different creep rupture envelopes 1 and 2 in Fig. 15.5. A time line through the intersection point between envelopes 1 and 2 will be used to establish the stress level for residual strength determination. At any time point on the (dashed) time line the Broutman–Sahu prediction (15.21) will give the same σ_R for Case 1 as for Case 2. It cannot distinguish creep rupture curve 1 from 2 in these cases, and for this reason its behavior is inconsistent and probably unacceptable. The kinetic crack method (15.22) does have the capability to distinguish Cases 1 and 2.

The Miner's rule and the Hashin–Rotem methods do not have the direct capability to make a prediction for $\tilde{\sigma}_R$.

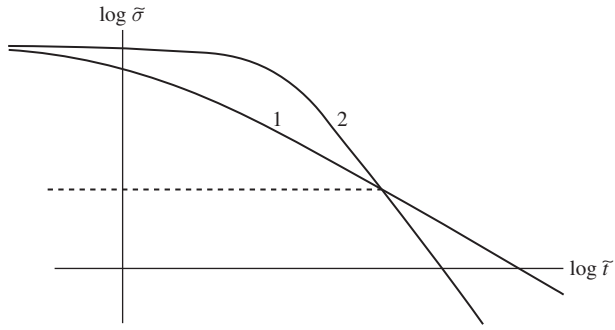


Fig. 15.5 *Residual strength for two cases.*

15.5 Life Prediction

Another useful way to compare the models is to consider the remaining life after a major damage event. To simulate this, a two-step loading program will be taken with the first step, of specified duration, being at a high stress overload, and the second step consisting of the service level stress up to failure.

The four damage models from (15.5), (15.6), (15.8), and (15.17) in this two-step form are given by

Miner's rule

$$\frac{\tilde{t}_1}{\tilde{t}_{c1}} + \frac{(\tilde{t} - \tilde{t}_1)}{\tilde{t}_{c2}} = 1 \tag{15.23}$$

Hashin–Rotem

$$\left(\frac{\tilde{t}_1}{\tilde{t}_{c1}}\right)^{\frac{\log \tilde{t}_{c2}}{\log \tilde{t}_{c1}}} + \frac{(\tilde{t} - \tilde{t}_1)}{\tilde{t}_{c2}} = 1 \tag{15.24}$$

Broutman–Sahu

$$\frac{\tilde{t}_1}{\left(\frac{1-\tilde{\sigma}_2}{1-\tilde{\sigma}_1}\right)\tilde{t}_{c1}} + \frac{(\tilde{t} - \tilde{t}_1)}{\tilde{t}_{c2}} = 1 \tag{15.25}$$

Kinetic-crack

$$\frac{\tilde{t}_1}{\left(\frac{\tilde{\sigma}_2}{\tilde{\sigma}_1}\right)^r \tilde{t}_{c2}} + \frac{(\tilde{t} - \tilde{t}_1)}{\tilde{t}_{c2}} = 1 \tag{15.26}$$

where index 1 refers to the first step and index 2 the second, and as before with

$$\tilde{t}_{c1} = \tilde{t}_c(\tilde{\sigma}_1)$$
$$\tilde{t}_{c2} = \tilde{t}_c(\tilde{\sigma}_2)$$

Before proceeding to the examples, a significant difference in the models can be seen from (15.23)–(15.26). In each case the first term represents the damage due to the stress overload, and the second term gives the normal accrual of damage at the service stress level, up to failure. The second terms in (15.23)–(15.26) are identical. Only the first terms—the major damage terms in the following examples—are fundamentally different. The first three forms have the first terms in them as mainly controlled by $\tilde{t}_{c1}$ but the last form, (15.26), has its first term as influenced by $\tilde{t}_{c2}$. The effective creep rupture time controlling the damage in the first step of (15.26) is as shown in Fig. 15.6.

From a damage growth point of view, for step 1 the damage grows up until time $\tilde{t}_1$, and that amount of damage is independent of whatever $\tilde{t}_{c1}$ may be. The kinetic crack model, (15.26), calibrates the first step damage relative to that which ultimately causes failure, $\tilde{t}_{c2}$, not to the hypothetical level given by $\tilde{t}_{c1}$.

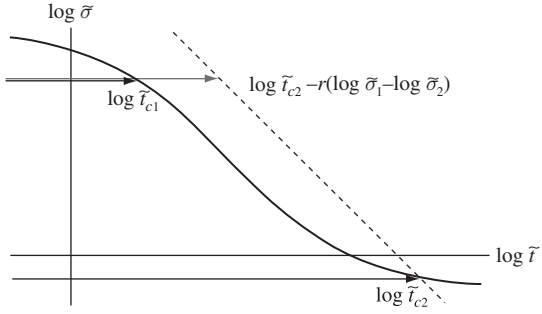


Fig. 15.6 *Effective creep-rupture time; first step of (15.26).*

The two damage/life examples to be given are specified by

$$\begin{aligned}
\tilde{\sigma}_1 &= 0.9 \\
\tilde{\sigma}_2 &= 0.5 \\
\tilde{t}_1 &= 9, \qquad \tilde{t}_{c1} = 10, \qquad \textit{Case A} \\
\tilde{t}_1 &= 252.1, \quad \tilde{t}_{c1} = 280.1, \quad \textit{Case B} \\
\tilde{t}_{c2} &= 10^5 \\
r &= 10
\end{aligned}$$

Only $\tilde{t}_1$ and $\tilde{t}_{c1}$ are varied between the two examples. Both cases have $\tilde{t}_1/\tilde{t}_{c1} = 0.9$. The governing creep rupture curves are shown schematically in Fig. 15.7. Case B represents that of a power-law behavior as specified by the power 10, which is in the range of polymer behavior.

Let the lifetime $\tilde{t}$ be specified by

$$\tilde{t} = \tilde{t}_1 + \Delta\tilde{t}$$

with $\Delta\tilde{t}$ designating the remaining life after the inflicted damage state of the first step. The results predicted by (15.23)–(15.26) are given in Table 15.1.

If there were negligible damage from the first step, the remaining life would be 10^5. Comparing the four models in Table 15.1 for Case A shows that the present model produces the least damage in the first step, while Miner's rule produces by far the most damage, resulting in the shortest

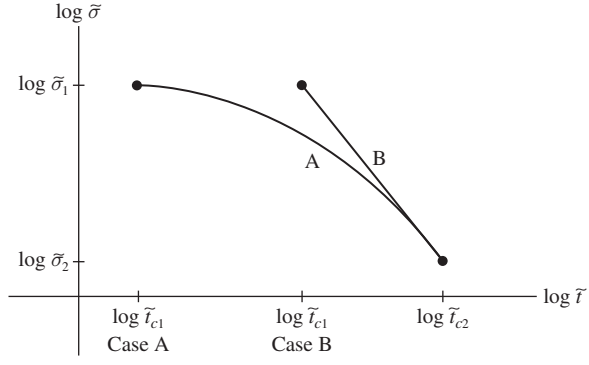


Fig. 15.7 *Creep-rupture curves for cases A and B.*

Table 15.1 Remaining life after damage

	Remaining Life, $\Delta\tilde{t}$			
	Kinetic-crack	Broutman–Sahu	Hashin–Rotem	Miner's rule
Case A	96,787.	82,000.	40,951.	10,000.
Case B	10,000.	82,000.	19,366.	10,000.

remaining life. There is no consistency between any of the models, as they all make completely distinct predictions in this case.

Comparing the four models for the power-law form, Case B, Broutman–Sahu produces a unrealistically small amount of damage (large remaining life), while the other three models are about the same at small values of remaining life.

As seen from Table 15.1, Miner's rule and Broutman–Sahu cannot distinguish the amount of stress overload damage between Cases A and B, even though the duration of the stress overload is 28 times greater in Case B than it is in Case A. It is only the kinetic crack model that shows properly the strong effect in going from Case A to Case B.

15.6 Residual Life

As the final evaluation condition, the problem of residual life will be posed. For a given program of stress history, if the load were terminated just an instant before failure, the question is: what would be the additional life at a reduced stress level? This is called the residual life problem Of course it would be extremely difficult to conduct this as an actual physical experiment because of the usual scatter in the data. However, it is very useful to use this as a conceptual test with various deterministic models to see which ones predict reasonable results and which ones predict unrealistic and unacceptable results.

The problem will now be posed in its simplest form as a two-step process. The duration of the first step is taken as $\tilde{t}_1 \rightarrow \tilde{t}_{c1}$, but understood to be terminated just before failure. The second step is then for $\tilde{\sigma}_2 < \tilde{\sigma}_1$ reduced by a specified amount. Since both $\tilde{t}_{c1}$ and $\tilde{t}_{c2}$ enter the problem formulation, it will be revealing to see which models involve both of these physical properties.

Let the residual life beyond step 1 be notated as

$$\tilde{t}_R = \tilde{t} - \tilde{t}_1$$

From (15.8) the Broutman–Sahu prediction for the residual life is given by

$$\tilde{t}_R = \left(\frac{\tilde{\sigma}_1 - \tilde{\sigma}_2}{1 - \tilde{\sigma}_2} \right) \tilde{t}_{c2} \tag{15.27}$$

From (15.17) the kinetic crack prediction for the residual life is given by

$$\tilde{t}_R = \tilde{t}_{c2} - \left(\frac{\tilde{\sigma}_1}{\tilde{\sigma}_2} \right)^r \tilde{t}_{c1} \tag{15.28}$$

Miner's rule (15.5) and Hashin–Rotem (15.6) both predict that there is no remaining residual life, even though the stress level is reduced. It is seen that only the kinetic-crack model brings both $\tilde{t}_{c1}$ and $\tilde{t}_{c2}$ into the residual life prediction.

As an example of residual life take

$$\begin{aligned} \tilde{\sigma}_1 &= 0.5 \\ \tilde{\sigma}_2 &= 0.4 \\ \tilde{t}_{c1} &= 10^5 \\ \tilde{t}_{c2} &= 9.5 \times 10^5 \\ r &= 10 \end{aligned}$$

The residual life predictions are

$$\begin{aligned} \tilde{t}_R &= 18{,}677. \quad \text{Kinetic-crack} \\ \tilde{t}_R &= 158{,}333. \quad \text{Broutman–Sahu} \end{aligned}$$

15.7 Conclusion

The three model testing conditions—damage/life, residual strength, and residual life—show that the kinetic-crack model is the only one of the four that satisfies these consistency tests. More specifically, the damage/life examples reveal the importance of the maximum slope of the creep

rupture-life envelope on log–log scales; Fig. 15.2. Only the kinetic-crack model includes this characteristic, as property r in (15.17). This property is important because it calibrates the kinetics of the flaw growth process. Without this defining property, two points on a stress-versus-life envelope cannot distinguish power-law behavior from anything else. It is analogous to trying to define curvature by only two points. With this property included it can be shown from (15.17) that in general for two-step programs of loading, the high-stress to low-stress sequence produces longer lifetimes than does the low-to-high sequence. Some observations confirm this specific sequence effect; Found and Quaresimin [15.12]. There are other data, however, that show just the opposite sequence effect. The one aspect on which there is general agreement is that Miner's rule shows no sequence effect and is inappropriate for general use.

Problem Areas for Study

1. Miner's rule for cumulative damage is a completely empirical form, yet it is shown in Section 15.3 that it produces the same results for life prediction as does kinetic-crack theory in the power-law range of behavior. This coincidence of results does not occur outside the power-law range. Is there an even more substantial physical basis to explain or justify Miner's rule?
2. What is the most important and most revealing conceivable test for the various cumulative-damage theories? Three tests were used herein, but there could be many more.
3. All the cumulative-damage theories in Chapter 15 presume states of proportional loading. How can one move beyond that effectively one-dimensional description toward a full tensor formalism?
4. It is common practice to introduce damage parameters as both scalars and tensors. The kinetic crack-growth method has no damage parameters. What are the trade-offs in having damage parameters versus not having them?
5. It is generally agreed that there would be expected to be a sequence effect in the order of application of different loading states, as cumulative damage is being built up. What are the physical reasons for this expected effect? Reason why there could be exceptions to this sequence effect.

References

[15.1] Nahshon, K. and Huchinson, J. W. (2008). "Modification of the Gurson Model for Shear Failure," *European Journal of Mechanics-A/Solids*, **27**, 1–17.

[15.2] Xue, L. (2008). "Constitutive Modeling of Void Shearing Effect in Ductile Fracture of Porous Materials," *Engineering Fracture Mechanics*, **75**, 3343–66.

[15.3] Suresh, S. (1998). *Fatigue of Materials*, 2nd ed., Cambridge University Press, Cambridge, UK.

[15.4] Krajcinovic, D. (1996). *Damage Mechanics*, Elsevier, North-Holland.

[15.5] Miner, M. A. (1945). "Cumulative Damage in Fatigue," *J. Applied Mechanics*, **12**, A159–64.

[15.6] Broutman, L. J. and Sahu, S. (1972). "A New Theory to Predict Cumulative Damage in Fiberglass Reinforced Composites," in *Composite Materials Testing and Design*, ASTM STP 497, 170–88.

[15.7] Hashin, Z., and Rotem, A. (1978). "A Cumulative Damage Theory of Fatigue Failure," *Mats. Sci and Eng.*, **34**, 147–60.

[15.8] Reifsnider, K. L. (1986). "The Critical Element Model: A Modeling Philosophy," *Eng. Frac. Mech.*, **25**, 739–49.

[15.9] Adam, T., Dickson, R. F., Jones, C. J., Reiter, H., and Harris, B. (1986). "A Power Law Fatigue Damage Model for Fibre-Reinforced Plastic Laminates," *Pro. Inst. Mech. Engrs.*, **200**, 155–65.

[15.10] Christensen, R. M. (2008). "A Physically Based Cumulative Damage Formalism," *Int. J. Fatigue*, **30**, 595–602.

[15.11] Post, N. L., Case, S. W., and Lesko, J. J. (2008). "Modeling the Variable Amplitude Fatigue of Composite Materials: A Review and Evaluation of the State of the Art for Spectrum Loading," *Int. J. Fatigue*, **30**, 2064–86.

[15.12] Found, M. S. and Quaresimin, M. (2003). "Two–Stage Fatigue Loading of Woven Carbon Fibre Reinforced Laminates," *Fatigue Fract. Engng. Mater. Struct.*, **26**, 17–26.

16
Probabilistic Failure and Probabilistic Life Prediction

This chapter involves the application of probabilistic methods to the prediction of failure for materials. Only the most severe and demanding of these types of problems are considered; namely, those of creep rupture and fatigue. Most other types of failure problems can be treated without the complications of statistical specification. Before explicitly investigating these "extreme" problems, some vital preliminaries must be developed. These concern not only the advantage but the necessity of using power-law forms to achieve generality in some situations, and also concern the most important probability distribution of all for materials applications: the famous Weibull distribution. A brief introduction poses the problems of interest.

16.1 Variability and Extreme Cases of Variability

All physical behaviors have variability, and this is especially true with materials. The amazing and extremely fortunate thing is how well deterministic theories capture most of what is happening. Nevertheless, there are some situations where the behaviors are so non-predictable by deterministic methods that statistical inference must be employed. Quantum mechanics is a prime example.

Usually when one averages over small scales to get behaviors at large scales the variability features blend into an effective continuum of average behaviors. This is the case with most macroscopic behaviors. In particular, for the failure of high-quality materials the static strengths are found to be tightly grouped around mean values that almost always suffice for predictions in applications. But there are exceptions, even at the macroscopic scale.

Failure conditions and specifications involve more than just the time-independent quasi-static strength. Time can be involved in explicit ways that are seemingly difficult to describe and understand at a simple level. Historically it was commonly observed that glass under low load levels could creep and ultimately even rupture. The lifetimes could be years or even centuries. The slow flow of glass is called creep, and the ensuing failure is called creep rupture. In modern terms, many materials that are normally thought to be completely inert and subject only to static failure actually undergo creep and creep rupture. All polymers fall under this category, as do metals at high temperature. The fascinating thing about this condition is that creep rupture in such materials can occur with little or no observable macroscopic creep preceding it. Obviously, processes of degradation are moving along at scales below those of macroscopic observation.

Completely similar situations occur with materials subjected to cyclic loading, which is the broad and important area of the fatigue behavior of materials. Virtually all materials types are susceptible to fatigue failure. So there is far more to failure than just static failure.

When one investigates these broader categories of failure, one finds that the common treatments and conditions that apply with static failure do not apply with these more general failure types. In particular, the tight grouping of failure data within a narrow band suddenly no longer seems to operate. The rule with these generalized failure types is that the test data are very widely scattered. The mean value of time to failure is almost meaningless, or at least misses the main aspect of behavior, the spread. It is one of the ironies of technical history that fatigue has nearly always been and usually still is represented as a single curve in an S–N plot that implies that mean values based on a few datapoints of fatigue life are perfectly acceptable. The reality is that typical fatigue data has so much scatter that it is obviously in need of a fully probabilistic treatment.

At this point the question arises as to when failure data can be treated using only mean values and when a full-blown probabilistic treatment must be employed. There is no simple and direct answer to this question, but the following can be said. Whenever logarithmic scales are needed to display meaningful datasets, then that is a quite clear indicator that it may be necessary to account for extreme variability of the material response. In particular, when involved with matters of creep-rupture lifetimes and fatigue-failure lifetimes, a probabilistic approach is certainly called for.

This chapter will examine failure problems that require the use of probability theory, and these will be mainly concerned with creep rupture and fatigue. Actually, it will be shown that both of these problem areas and methods of solution devolve from the same formalism. One is measured explicitly in terms of elapsed time, while the other is in number of cycles, which is an equivalent scale for considering the accumulation of damage ultimately culminating in failure.

Although many different probability functions have been applied to these classes of problems, the attention here will be focused upon only one of these: the Weibull distribution. From a mathematical point of view the Weibull distribution is just one of a great many different probability distribution functions, but from a materials point of view it has a very special status. Its special meaning and applicability will be fully explored here. The probability function of the Weibull type will be derived here in a manner that naturally displays its significance, and thereafter it will be applied to the life-prediction problems of interest.

The entry to the lifetime-prediction problem is given by the time-controlled growth of flaws and defects. A theoretical approach, appropriately named kinetic crack growth theory, will be developed and applied to the lifetime problem, first in deterministic form, then with generalization to probabilistic forms.

Before embarking on this life-prediction course an even more basic concept must be developed and assimilated. This concept is that of the power-law form, widely employed in physics as representations for various effects linking descriptive variables. Power laws for specific applications to materials problems will be closely examined and developed. They are not merely empirical forms, attractive only for their ease of use. In the present materials context they will be found to be indispensible. It is thus necessary to start with the development and basis for the power-law form in materials failure applications, even before deriving the Weibull distribution.

16.2 Power-Law Failure Interpretation

Power-law behavior is commonly seen with the non-destructive properties of polymers; namely, relaxation functions and creep functions. This account is begun by seeking to understand the source for the power-law behavior with the failure of polymers.

When dealing with polymeric materials behavior, the relaxation functions and creep compliance functions are often represented through their spectra; Ferry [16.1]. It is equally well motivated to represent creep rupture (failure) behavior through spectra, using the same formalism; thus take

$$\sigma(t_c) = \int_0^\infty \tilde{H}(\tau) e^{-\frac{t_c}{\tau}} d\tau \tag{16.1}$$

where the constant stress causing failure at time t_c is σ, and $\tilde{H}(\tau)$ is the spectrum.

For a single mechanism of failure there would be a single exponential describing the failure-time form,

$$\sigma(t_c) = A e^{-\frac{t_c}{\tau_1}} \tag{16.2}$$

The spectrum is then that of a single δ-function at $\tau = \tau_1$. The simple form (16.2) functions very poorly in representing actual creep rupture behavior. This in itself produces a lead on how to search for a more general form.

Before proceeding further it is helpful to recognize the advantage of using log scales, and following Ferry [16.1] take (16.1) in the alternative form

$$\sigma(t_c) = \int_0^\infty \hat{H}(\tau) e^{-\frac{t_c}{\tau}} d(\log \tau) \tag{16.3}$$

where

$$\hat{H}(\tau) = \tau \tilde{H}(\tau) \tag{16.4}$$

and where the actual creep rupture failure process is now recognized to involve a continuum of changing failure modes, not just a single event.

The single δ-function type of spectrum is singularly different from the function it attempts to represent, $\sigma(\tau_c)$. This single δ-function is one limiting case, and it works poorly. One could add more δ-functions spaced at different increments of time, but instead it is most helpful to go to the opposite limiting case by taking the spectrum and the function it represents as being identical, to within a scaling constant. This will produce the most regular and continuously distributed spectrum.

Following this course, take

$$\hat{H}(\tau) = \alpha\sigma(t_c)|_{t_c=\tau} \tag{16.5}$$

where α is some constant.

At this juncture there is a useful approximation to the spectra forms in (16.1) and (16.3). This is obtained by approximating the exponential in (16.1) by a step function that is positioned such that the total square error between the two is minimized; see Christensen [16.2]. Then the resulting form can be differentiated to obtain

$$\tilde{H}(\tau) \tilde{=} - \frac{d\sigma(\tau)}{d\tau} \tag{16.6}$$

with t_c replaced by τ. A time scaling factor of Ln 2 = 0.693 is involved in this approximation but it is not needed here. This approximation is valid for large values of τ but not for small τs. For the logarithmic scale case, then

$$\hat{H}(\tau) \tilde{=} - \tau \frac{d\sigma(\tau)}{d\tau} \tag{16.7}$$

Next substitute (16.5) into (16.7) to obtain

$$\frac{d\sigma(\tau)}{d\tau} + \alpha \frac{\sigma(\tau)}{\tau} = 0 \tag{16.8}$$

The solution of (16.8) is

$$\sigma(\tau) = B\tau^{-\alpha}$$

or

$$\sigma(t_c) = Bt_c^{-\alpha} \tag{16.9}$$

Thus the power law form is the limiting case where the function and its spectrum are related by (16.5).

The power law form (16.9) has the best possibility to model physical behavior over many decades of time. The power law form is not just an

empirical term used only for its convenience in manipulations. It has a physical basis as the smoothest, most regular descriptive form for distributed failure mechanisms activated by stress. Power law forms arise in a variety of situations involving materials failure. Ample use—even crucial use—will be made of power law forms in the following developments. In fact, it will be found to provide the "spine" for the Weibull distribution.

16.3 Weibull Distribution Physical Basis

The Weibull distribution for use in probability theory is of surprisingly recent formalization and implementation; Weibull [16.3], 1951. This seminal paper is immensely readable, enjoyable, and convincing in its application to many different systems and situations. A general and complete treatment of the subject is now available and is presented by Rinne [16.4], including historical antecedents. The validity of the Weibull distribution for any of a wide variety of applications usually follows from the standard probabilistic approach: if it fits the data then it is used, but if it does not then some other distribution function must be tried. It does work surprisingly well in a great many cases.

The explicit justification for the use of the Weibull distribution with materials failure problems usually appeals to the weakest-link argument involving the weakest link (probabilistically) in a chain of links. It can give one pause to read in some books that this extremely simple one-dimensional pattern of a chain of links provides the supporting foundation for the use of the Weibull distribution with the failure behavior of three-dimensional continua; namely, all homogeneous materials. Its applicability and success in the technical area is not in question, but what is missing is a substantial basis or development of why it is so successful. This seeming paradox will be addressed here to help solidify its use with materials. The investigation is at the macroscopic scale of most applications.

First, some appropriate terminology must be established. Some of the examples will use time as the independent variable. A collection of materials samples under load of some type will fail in a sequential manner as time elapses. This could be the service life of electric lights or the function of turbine blades, or just anything that does not have absolutely perfect predictability. Whatever it is, serving in whatever environment, the cessation of function is defined as failure, and the items of interest are taken as a random variable of time. The times of failure can be gathered

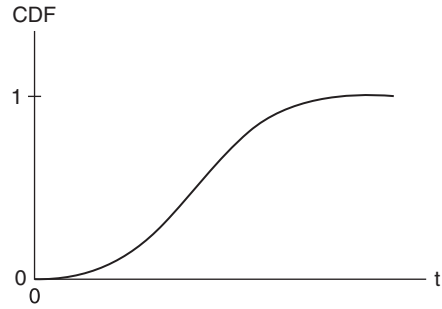


Fig. 16.1 *Cumulative distribution function.*

together, and the total number of failures up to any specific time are known. The collection of data suggests the existence of a cumulative distribution function (CDF) (of failure), as shown in Fig. 16.1.

The derivative of the CDF with respect to time forms the probability density function (PDF). It shows the band of the dominant failure occurrences, and often falls off on either side of its peak. The complement of the cumulative distribution function is given by (1 – CDF), and it provides the time record of the probability of survivors. Time could be replaced by any other variable of interest. Stress as well as time, of course, will enter all matters here.

It is tremendously advantageous to characterize these forms by analytical functions, the normal distribution being the most common PDF and CDF. Such distributions can then be used to predict a variety of effects. Predictability is the operative term. Even though the absolute time to failure cannot be determined with certainty, the probability of failure and of survival can be expressed quantitatively and used with complete assurance.

In addition to the PDF, CDF, and 1 – CDF, there is a fourth basic function that will be of relevance here. From reliability theory the hazard rate or hazard function is defined by

$$h = \frac{\dfrac{d}{dt}(CDF)}{1 - CDF} \tag{16.10}$$

The hazard function is defined as the instantaneous rate of failure, since the denominator in (16.10) quantifies the survivors at the value of time t.

To avoid confusion with the terminology of the instantaneous static strength, the hazard function will be referred to here simply as the rate of failure, rather than the instantaneous rate of failure. The term "rate of failure" applies whether it is the rate with respect to time or with respect to increase in stress, or anything else that is changing in a controlled manner.

With these terms the derivation of a CDF for the failure of materials can begin. Rather than starting with the time-dependent failure that occurs in creep rupture (or fatigue), it is helpful to start with the simplest case: namely, the static failure of materials. The complications of time-dependent creep rupture will be treated later. Time is not explicitly involved in determining the static strength of materials, since it is normally done quickly compared with the time scales of interest in creep rupture. Accordingly, this will be called the instantaneous static strength, or simply the instantaneous strength, or even just the static strength, and will be designated by σ_i. Normally this is taken to be a single scalar representing the mean value from a few tests of specimens. Here, however, it is advisable to pay full attention to the variability in the instantaneous strength, because this may provide a clue or lead to the more serious effects that will follow in the time domain. The hazard rate, rate of failure, for the instantaneous static strength is then

$$h = \frac{\dfrac{d}{d\sigma}(CDF)}{1 - CDF} \tag{16.11}$$

The objective now is to determine the CDF for the instantaneous static strength of materials by any means that recognizes and respects the characteristics and capabilities of modern high-quality materials, whether they be metals, polymers, ceramics, glasses, or anything else.

Only the most regular forms are expected to be likely to apply to these engineering materials. This smoothness or regularity of behavior is consistent with that for many (but not all) physical systems. Furthermore, it is far easier and more direct to specify realistic forms for h, the rate of failure function, than it is to conjecture forms for the PDF or CDF. This is because the hazard rate or (probabilistic) rate of failure has a physical meaning that can be grasped intuitively, as will be seen.

There is no method to identify a unique form for the rate of failure, h, of the static strength, but having just seen in the preceding section

the "power" of the power law representation, it is not only reasonable but compellingly logical to start with the power law form for h. Let h be given by

$$h = A\sigma^p \tag{16.12}$$

where p is the power law exponent. The allowable range for p for the instantaneous strength is not yet apparent, but it is seen that (16.12) embodies several different possibilities, as shown in Fig. 16.2.

The three cases shown in Fig. 16.2 involve (i) an increasing rate of failure with an accelerating rate of failure, (ii) an increasing rate of failure but with a decelerating rate of increase, and (iii) a decreasing rate of failure. All three distinctly different behaviors are controlled and specified by the single parameter, p. Relation (16.12) also admits special behaviors and interpretations at $p = 0$ and $p = 1$, as will be seen.

The behavior shown in Fig. 16.2 for the power law form of h provides the physical basis for what will turn out to be the Weibull distribution. The probabilistic rate of failure (16.12) as a function of σ stress (and stress change) is a monotonically changing function of σ with a monotonically changing first derivative. This is by far the most regular and most likely physical occurrence for the static strength behavior of engineering materials of widespread application. The Weibull distribution is not the

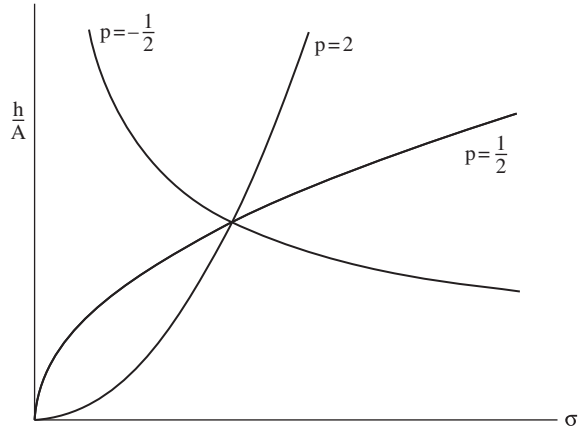


Fig. 16.2 *Hazard (rate of failure) function for (16.12).*

only one with a monotonic rate of failure form, but it does appear to be by far the most versatile and most general two-parameter form of all of them.

For a rate of failure that decreases with increasing stress, surely the resulting distributions would be broad and diffuse. Conversely, for a rate of failure that increases with increasing stress the distributions would seem to be necessarily narrow and concentrated. Exactly these physical effects will be found to occur with the CDF that results from the power law form for h, (16.12). Intuitively, it would be expected that most engineering materials that are sufficiently highly developed to admit the characterization of being homogenous (uniform at the scale of intended use) would fall within category (i) above, having an increasing probabilistic rate of failure and with acceleration with respect to increasing stress. Interestingly, in contrast to the static strength, lifetimes will be found to conform to the decreasing rate of failure form as time increases.

Combining (16.11) and (16.12) gives

$$\frac{\frac{d}{d\sigma}(CDF)}{1-CDF} = A\sigma^p \tag{16.13}$$

The solution of this differential equation is given by

$$CDF = 1 - e^{-\left(\frac{\sigma}{\sigma_s}\right)^{p+1}} \tag{16.14}$$

where

$$\sigma_s = \left(\frac{p+1}{A}\right)^{\frac{1}{p+1}} \tag{16.15}$$

Finally, let

$$p + 1 = m \tag{16.16}$$

which then gives (16.14) as

$$CDF = F(\sigma) = 1 - e^{-\left(\frac{\sigma}{\sigma_s}\right)^m} \tag{16.17}$$

where the power law form (16.12) from which this is derived has

$$A = \frac{m}{\sigma_s^m} \tag{16.18}$$

$$p = m - 1 \tag{16.19}$$

The CDF in (16.17) is that of the two-parameter Weibull distribution with

$$m - \textit{shape parameter}$$

$$\sigma_s - \textit{scale parameter}$$

Large values for m give a very narrow, concentrated PDF, while small values give very broad distributions. The values $m = 1$ ($p = 0$) is that for the exponential distribution, while $m = 2$ ($p = 1$) gives the Rayleigh distribution. The shape parameter must have $m > 0$, so then the associated hazard rate power law form (16.12) has $p > -1$.

Thus the Weibull distribution is derived from the most fundamental form for the hazard, rate of failure function. The power law form (16.12) for the rate of failure has a very wide range of realistic physical behaviors, and consequently and subsequently provides the very powerful formalism of the Weibull distribution. The chain-of-links scenario was of no use or relevance in this derivation.

16.4 Kinetic Crack Theory and Life Prediction

The long time behavior of materials in creep and fatigue conditions is very important for a wide range of applications. Unfortunately these are some of the most difficult problems that can be confronted insofar as the basic mechanisms of damage and failure are concerned. Nevertheless, it is necessary to treat these problems, especially since there is so much scatter in typical testing data for these problems.

The ultimate goal is to obtain a full and complete probabilistic treatment for these problems, but the initial approach will be to derive deterministic forms for the subject of life prediction, mainly aimed for creep rupture in polymers and metals at high temperatures, and for fatigue in all materials types. The approach uses the theory of kinetic

crack growth, as small-scale damage, and generally follows Christensen and Miyano [16.5], which in turn follows much earlier work. The first part also has some overlap with the cumulative damage treatment in Chapter 15. The notation is changed slightly from that in Chapter 15 to accommodate probabilistic variables.

Take an elastic material as having an initial state of flaws, here idealized as that of the central crack problem under Mode I conditions with the initial crack size of $2a_0$. Rapidly applied loads will cause instability of the crack due to fracture. The stress at which this occurs is taken as σ_i, here referred to as the instantaneous or static strength of the material. For stress levels lower than σ_i, controlled crack growth is taken to occur. The crack growth will continue up to the time at which the crack becomes sufficiently large such that at the then existing stress level, fracture will occur, this being by the same basic mechanism that causes the fracture instability at σ_i. The central problem is to determine this time to failure under a given stress history.

Once again it is necessary to appeal to and rely upon the power law representation in order to proceed further. The crack growth rate is taken to be controlled by the kinetic power law form

$$\dot{a} = \frac{da}{dt} = \lambda(\sigma\sqrt{a})^r \tag{16.20}$$

expressed in terms of the stress intensity factor of the central crack problem from classical fracture mechanics. The crack size and far-field stress in (16.20) have general time dependence, and where in (16.20) r is the power law exponent and λ is a material parameter. Equation (16.20) is written in the form shown because the central crack problem is being considered here. The power law form (16.20) for the time-dependent crack growth problem is sometimes called the Paris law in the context of cyclic fatigue. The advantages and significance of power law forms have been discussed already.

Separate the variables in (16.20) and formally integrate to obtain

$$\int_{a_0}^{a(t)} \frac{da}{a^{\frac{r}{2}}} = \lambda \int_0^t \sigma^r(\tau)\, d\tau \tag{16.21}$$

where a_0 is the initial size of the crack. Performing the explicit integration gives

$$\left(\frac{a}{a_0}\right)^{-\frac{r}{2}+1} - 1 = \lambda\left(1 - \frac{r}{2}\right) a_0^{\frac{r}{2}-1} \int_0^t \sigma^r(\tau)\, d\tau \tag{16.22}$$

Relation (16.22) gives the crack size $a(t)$ as a function of time, for a given stress history. In effect, the history of the stress intensity factor is now considered as known. The time flow in (16.22) will be allowed to continue up to the point at which the crack becomes unstable, which will be considered next.

Now, at the end of the lifetime, instantaneous, unstable failure occurs. At this time of failure, $t = t_f$, the fracture condition is specified by the critical value of the stress intensity factor, thus

$$\left[\sqrt{a(t)}\sigma(t)\right]_{t=t_f} = \sqrt{a_0}\sigma_i \tag{16.23}$$

Relation (16.23) provides the key to the present approach. The right-hand side of (16.23) is the critical stress intensity factor for the instantaneous static strength, while the left-hand side is necessarily the same critical stress intensity factor but at the end of the lifetime under a load less than σ_i but a crack size greater than a_0. At $t = t_f$ rewrite (16.23) as

$$\frac{a}{a_0} = \left(\frac{\sigma_i}{\sigma}\right)^2 \tag{16.24}$$

Substitute (16.24) into (16.22) to obtain

$$\left(\frac{\sigma_i}{\sigma}\right)^{2-r} - 1 = \lambda\left(1 - \frac{r}{2}\right) a_0^{\frac{r}{2}-1} \int_0^{t_f} \sigma^r(\tau)\, d\tau \tag{16.25}$$

Relation (16.25) can be put into non-dimensional form. Let stress be non-dimensionalized by the instantaneous static strength as

$$\tilde{\sigma} = \frac{\sigma}{\sigma_i} \tag{16.26}$$

and let time be non-dimensionalized as

$$\tilde{t} = \frac{t}{t_1} \tag{16.27}$$

where

$$t_1 = \frac{a_0^{1-\frac{r}{2}} \sigma_i^{-r}}{\lambda\left(\frac{r}{2} - 1\right)} \tag{16.28}$$

With (16.26)–(16.28) the lifetime relation (16.25) becomes

$$1 - \tilde{\sigma}^{r-2}(\tilde{t}_f) = \int_0^{\tilde{t}_f} \tilde{\sigma}^r(\tau) d\tau \tag{16.29}$$

Finally, drop the $\tilde{t}_f$ notation in (16.29), where $\tilde{t}$ will be understood to be the lifetime, and then (16.29) is

$$\frac{1}{1 - \tilde{\sigma}^{r-2}(\tilde{t})} \int_0^{\tilde{t}} \tilde{\sigma}^r(\tau) d\tau = 1 \tag{16.30}$$

This is the final form of the general lifetime criterion. Specific cases follow from (16.30) when the stress history is specified.

For a given stress history, $\tilde{\sigma}(\tau)$, the form (16.30) determines the lifetime $\tilde{t}$, to failure. The parameter t_1 in the non-dimensionalized time (16.27) will be treated as a single free parameter, rather than using (16.28). The time parameter t_1 is effectively substituted for the parameter λ in (16.28). Parameter t_1 is accommodated by shifts along the log time axis. Relation (16.30) contains only two parameters—the power law exponent, r, and the time shift parameter, t_1—considering the instantaneous static strength, σ_i, to be known.

Consider the creep rupture condition that is specified by constant stress. From (16.30) the lifetime is found to be

$$\tilde{t}_c = \frac{1}{\tilde{\sigma}^r} - \frac{1}{\tilde{\sigma}^2} \tag{16.31}$$

The creep rupture lifetime result is shown schematically in Fig. 16.3.

The short time range is that of the instantaneous static strength asymptote, and the long time range is the power law asymptote. The comparison of the creep rupture theoretical prediction with some experimental data is shown in Fig. 16.4.

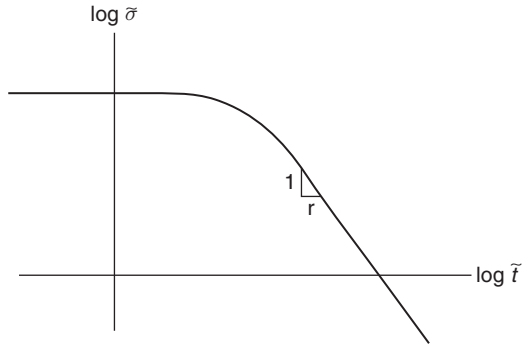


Fig. 16.3 *Creep rupture.*

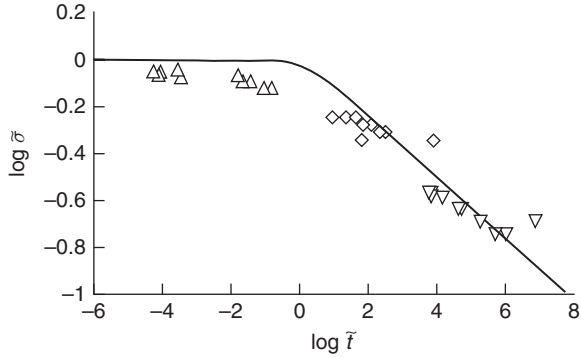


Fig. 16.4 *Creep rupture (16.31) and data.*

The data is for a carbon fiber–vinyl ester resin laminate. The data is from Christensen and Miyano [16.6] which includes references to the earlier experimental work. It is seen that the data round the fairly sharp corner of the theoretical prediction. This would probably occur in all applications to complex systems. An empirical form is given in [16.6] that fits the data quite closely.

As an example of deterministic life prediction for a non-constant loading history, consider the case of constant strain rate. Since the bulk material is perfectly elastic, the case of constant strain rate is the same as constant stress rate then specified by

$$\sigma(\tau) = \beta\tau \tag{16.32}$$

Write the lifetime criterion (16.30) in dimensional form using (16.26) and (16.27) as

$$\frac{1}{\sigma_i^2 t_1 \left[\sigma_i^{r-2} - \sigma^{r-2}(t)\right]} \int_0^t \sigma^r(\tau) d\tau = 1 \tag{16.33}$$

Substituting (16.32) into (16.33) and carrying out the integration gives

$$\beta^r t^{r+1} + (r+1)\beta^{r-2}\sigma_i^2 t_1 t^{r-2} - (r+1)t_1 \sigma_i^r = 0 \tag{16.34}$$

In general this is a high-order polynomial to be solved for the time to failure, t. Rather than doing that directly, it is advantageous to eliminate the stress rate value β in favor of the stress at failure. Using (16.32), write β as

$$\beta = \frac{\sigma_f}{t_f} \tag{16.35}$$

where σ_f is the stress at the failure time $t = t_f$. With (16.35) then (16.34) becomes

$$\sigma_f^r t_f + (r+1)\sigma_i^2 t_1 \sigma_f^{r-2} - (r+1)t_1 \sigma_i^r = 0 \tag{16.36}$$

This relation directly gives the time to failure as

$$\frac{t_f}{t_1} = (r+1)\left[\left(\frac{\sigma_i}{\sigma_f}\right)^r - \left(\frac{\sigma_i}{\sigma_f}\right)^2\right] \tag{16.37}$$

Take the logarithm of (16.37) to find

$$\log\frac{t_f}{t_1} = \log\left[\left(\frac{\sigma_i}{\sigma_f}\right)^r - \left(\frac{\sigma_i}{\sigma_f}\right)^2\right] + \log(r+1) \tag{16.38}$$

From (16.38) it is seen that the constant strain rate lifetime is the same as the creep rupture result (16.31) but shifted by the amount

$$\log(r+1)$$

Along the log time scale this shifting property is shown in Fig. 16.5.

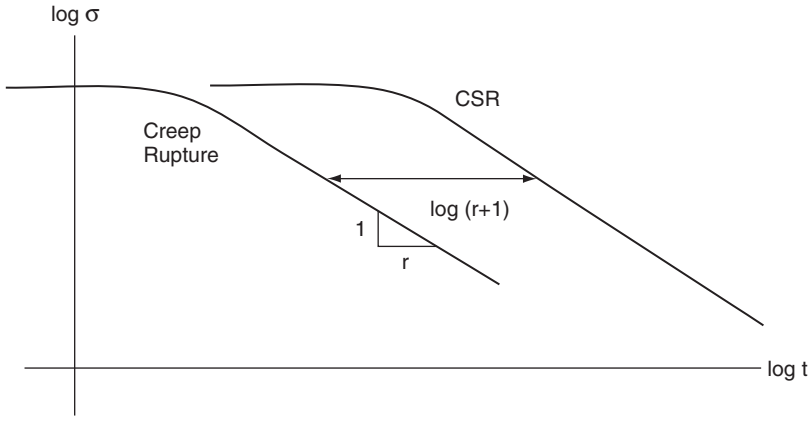


Fig. 16.5 *Constant strain rate shifting property.*

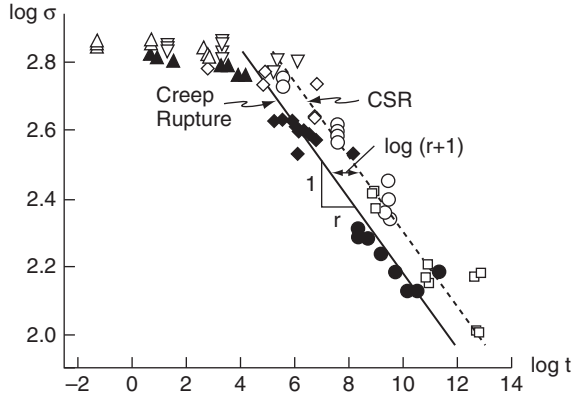


Fig. 16.6 *Constant strain rate prediction and data.*

The shifting property has been experimentally verified using data from a polymeric fiber composite laminate, Fig. 16.6, from [16.5].

16.5 Probabilistic Generalization of Strength and Lifetime Theory

When one first sees the raw data from creep rupture testing, the problem seems hopeless. There is no apparent or discernable order to the extremely scattered data. Yet when the proper probabilistic treatment is applied, order does indeed begin to emerge.

Now, from the previous deterministic formulation let the instantaneous static strength, σ_i, be characterized as a random variable with a specific distribution function. This physical effect results from the inherent scatter in the initial flaw sizes or weaknesses in the virgin material.

Let deterministic σ_i be specified through a Weibull distribution where the complement of the CDF, $F(\sigma)$, is given by

$$F(\sigma) = e^{-\left(\frac{\sigma}{\sigma_s}\right)^m} \tag{16.39}$$

where m is the shape parameter and σ_s is the scale parameter. Now let

$$F(\sigma) = 1 - k \tag{16.40}$$

where k is the quantile of failure for the instantaneous static strength, with $k = 0$ indicating no failures. It follows that

$$\sigma_k = \sigma_s[-Ln(1-k)]^{1/m} \tag{16.41}$$

where now σ_k is the probabilistic form for σ_i.

The deterministic creep rupture result (16.31) is rewritten here as

$$\tilde{t} = \left(\frac{\sigma_i}{\sigma}\right)^r - \left(\frac{\sigma_i}{\sigma}\right)^2 \tag{16.42}$$

Using (16.41) the probabilistic generalization of (16.42) is given by

$$\tilde{t} = \frac{[-Ln(1-k)]^{\frac{r}{m}}}{(\frac{\sigma}{\sigma_s})^r} - \frac{[-Ln(1-k)]^{\frac{2}{m}}}{(\frac{\sigma}{\sigma_s})^2} \tag{16.43}$$

Write (16.43) in symbolic form as

$$\tilde{t} = \hat{f}\left(\frac{\sigma}{\hat{\phi}(k)}\right) \tag{16.44}$$

where

$$\hat{\phi}(k) = \sigma_s[-Ln(1-k)]^{1/m} \tag{16.45}$$

and

$$\hat{f}(x) = \frac{1}{x^r} - \frac{1}{x^2} \tag{16.46}$$

The results (16.44)–(16.46) give the probabilistic time to failure for the creep rupture condition as a function of the Weibull shape and scale parameters of the instantaneous static strength, the power law exponent, r, and the specified quantile of failure, k. The same probabilistic generalization can be applied to the lifetime form (16.30) for any prescribed stress history. In an important application, the kinetic crack formulation will be shown to account for the time dependent, progressive debonding of statistically imperfect fibers until the associated composite can no longer sustain the total load and probabilistic creep rupture occurs.

Now examine the power law range of behavior for creep rupture. From (16.43) this range is specified by

$$\tilde{t} = \left(\frac{\sigma_s}{\sigma}\right)^r [-Ln(1-k)]^{r/m} \tag{16.47}$$

Solve (16.47) for $(1 - k)$ to obtain the complement to the lifetime CDF as

$$1 - k = e^{-\left[\frac{t}{t_1\left(\frac{\sigma_s}{\sigma}\right)^r}\right]^{\frac{m}{r}}} \tag{16.48}$$

From (16.48) it is seen that the probabilistic lifetime in the power law range is Weibull distributed with

$$\frac{m}{r} - \textit{lifetime shape parameter} \tag{16.49}$$

$$t_1\left(\frac{\sigma_s}{\sigma}\right)^r - \textit{lifetime scale parameter} \tag{16.50}$$

The results (16.49) and (16.50) not only are unusually compact but extraordinarily comprehensive. The lifetime probabilistic behavior is completely specified by the instantaneous static strength Weibull parameters m and σ_s and the slope of the lifetime envelopes in the power law range, $1/r$. Parameter t_1 is most easily found directly from lifetime data. The significance of these results will be examined next.

In the 1970s, '80s, and '90s a nearly unique program of creep rupture testing was conducted at Lawrence Livermore National Laboratory. T. T. Chiao formulated and managed a very large program of testing involving many hundreds of specimens. Some of them were maintained for as long as many years under load before failure. The testing specimens were thin strands of unidirectional fibers impregnated with epoxy resins. The three fiber types were carbon, aramid, and glass. Dead loads were applied at different loading levels, and timing devices recorded failure. It was a meticulously planned program with dedicated long-term follow-through. The testing results are known as the LLNL databases for creep rupture.

These creep rupture databases were used to evaluate this highly idealized probabilistic lifetime theory. In particular, using the result (16.49) the predicted lifetime shape parameters were compared with the measured lifetime shape parameters. There could hardly be a more critical evaluation than this. The lifetime failure data are spread over many decades, while the instantaneous static strength data are tightly grouped around a mean stress level. The evaluation thus tests both the Weibull distribution hypothesis for static strength and lifetime, as well as the kinetic crack life prediction theory. All of the datasets were found to be satisfactorily modeled by Weibull distributions, and the shape and scale parameters were determined by the method of maximum likelihood. All exhibited the expected power law ranges of behavior.

The especially important case of AS-4 carbon-epoxy was analyzed by Christensen and Glaser [16.7]. The Kevlar aramid–epoxy system was also analyzed by Christensen and Glaser [16.8]. Finally, the S glass–epoxy system was analyzed by Glaser, Christensen, and Chiao [16.9]. The basic theory was developed in the 1970s and '80s by Christensen, in other publications that were the forerunners of much of what is presented in this section and the previous one. The notation used in all these references is a little different from that used here, which is simpler, but they all are compatible.

The probabilistic data reductions and evaluations directly from the above three references are summarized in Table 16.1. The term "measured" means determined from the databases, while "predicted" refers to the theoretical result (16.49) coming from (16.48) for the shape parameter of lifetime.

Thus the enormously broad distributions of the creep rupture lifetimes are successfully predicted by the extremely narrow distributions of

Table 16.1 Creep rupture data and evaluation of theory

Carbon–epoxy		
Power law exponent r	72.2	Measured
Static strength shape parameter	16.0	Measured
Lifetime shape parameter	0.191	***Measured***
Lifetime shape parameter	0.222	***Predicted***
Aramid–epoxy		
Power law exponent r	45.6	Measured
Static strength shape parameter	46.3	Measured
Lifetime shape parameter	0.960	***Measured***
Lifetime shape parameter	1.02	***Predicted***
Glass–epoxy		
Power law exponent r	32.8	Measured
Static strength shape parameter	30.6	Measured
Lifetime shape parameter	0.852	***Measured***
Lifetime shape parameter	0.933	***Predicted***

the static strength and the slopes of the probability quantiles in the lifetime power law ranges. It is now easy to see how designing for lifetime probabilities becomes a practical proposition using only relatively few specimens for testing. Considering the importance and complexity of the problem, the clarity of this solution is something quite special. Unifying results such as these do not come along very often in such difficult fields.

Now the Weibull specific and power law specific results (16.44)–(16.46) will be generalized to any corresponding form by taking the non-dimensional creep rupture time to failure $\tilde{t}$ as

$$\tilde{t} = f\left(\frac{\sigma}{\phi(k)}\right) \tag{16.51}$$

where

$$\phi(k) - \textit{not necessarily of Weibull form}$$

and

$$f\left(\frac{\sigma}{\phi(k)}\right) - \textit{not necessarily of power law form}$$

The Weibull and power law forms of $\varphi(k)$ and $f(x)$ are given by (16.45) and (16.46), but now for the more general forms all that is required is

that $\varphi(k)$ be any form controlling the probabilistic instantaneous static strength, and σ and $\varphi(k)$ enter $f(x)$ only in the combination $\mathrm{f}\left(\frac{\sigma}{\phi(k)}\right)$ and that $\mathrm{f}\left(\frac{\sigma}{\phi(k)}\right) = 0$ when $\frac{\sigma}{\phi(k)} = 1$; compare with (16.46). Thus the forms of $\varphi(k)$ and $f(x)$ remain quite general.

Express the form (16.51) in terms of log variables. This gives (16.51) as

$$\log \tilde{t} = g\left(\log\left(\frac{\sigma}{\varphi(k)}\right)\right) \tag{16.52}$$

where the function $g(\log\ x)$ is found from the form of $f(x)$ in (16.51). Rewrite (16.52) as

$$\log \tilde{t} = g(\log(\sigma) - \log\phi(k)) \tag{16.53}$$

The result (16.53) on $\log\ \sigma$ versus $\log\ t$ scales has the form shown in Fig. 16.7 for different values of the quantile of failure, k.

The result (16.53), as shown in Fig. 16.7, reveals that the different probability levels of failure have the form of a single, master curve that is shifted vertically along the $\log\ \sigma$ axis. This unexpected and extremely simplifying type of behavior was first deduced by Christensen [16.10].

Relative to the vertical shifting behavior shown in Fig. 16.7, twenty specimens of unidirectional carbon fiber–vinyl ester composites were tested in the constant strain rate condition at elevated temperature to simulate the time-dependent case, and were also tested for instantaneous static strength. The results are shown in Fig. 16.8; Christensen and Miyano [16.5].

The vertical shift result was evaluated as shown. The results were found to be of Weibull distribution, Fig. 16.8, with the two-shape parameters of

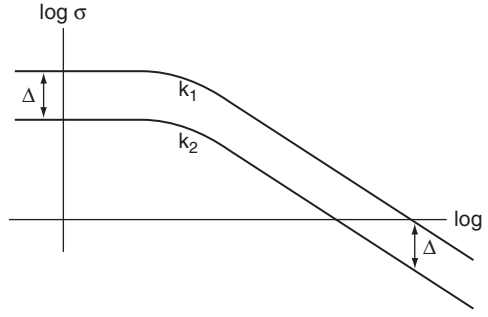
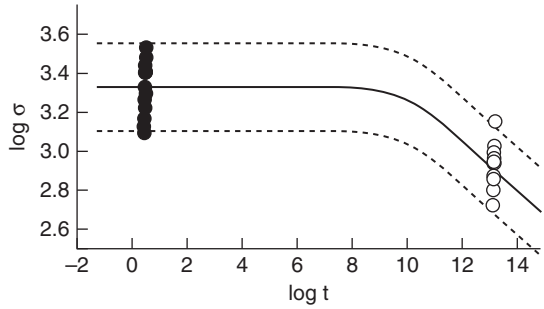
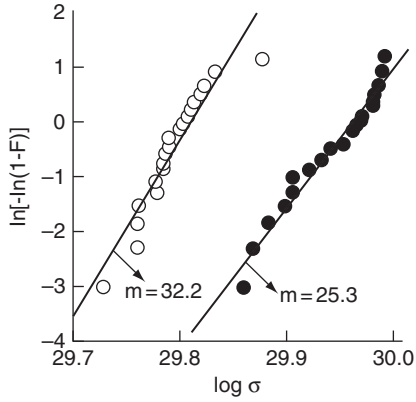


Fig. 16.7 *Vertical shifting property.*



(a) Data of tensile CSR



(b) Weibull probability

Fig. 16.8 *Weibull probability for vertical shifting.*

$m = 32.2$ and $m = 25.3$. While these are not quite as close as would be ideal, they do show consistency with the vertical shift behavior. That they are both of the Weibull type is also consistent with the vertical shift.

This completes the coverage of primarily using the Weibull distribution with the cases of probabilistic failure. Successful though it is, the Weibull distribution is not the answer to all problems in the macroscopic domain. Sometimes a particular Weibull distribution satisfactorily covers a certain range, but then another distribution seems to explain the more extreme ranges and cases. This secondary distribution could be a different Weibull distribution such as the exponential distribution, or it could be an entirely different type of distribution. Even in cases such as the later case, the Weibull distribution still seems to cover the main aspects of the probabilistic failure behavior.

Needless to say, this has been, and continues to be, an active field. There is a broad and important literature in the field in addition to the book by Rinne [16.4] and many other books, as well as the references already given. Bazant and colleagues [16.11]–[16.13] have provided a comprehensive treatment of probabilistic failure based upon theories of behavior at the nanoscale. Phoenix and colleagues [16.14]–[16.16] have provided longstanding treatments of probabilistic failure, primarily for composites. Miyano, Nakada, and colleagues [16.17]–[16.19] have provided a very extensive literature of testing results, including probabilistic behavior obtained by the method of accelerated testing using elevated temperatures for polymeric-based composite fiber systems.

The problem of cyclic fatigue in metals admits an analogy with the present problem of creep rupture in polymers. Taking (16.20) as da/dn rather than da/dt, with n, the number of cycles, as the measure of duration, then all of the statistical forms and conclusions found here for creep rupture have a one-to-one counterpart in the cyclic fatigue area. Even though the controlling properties (power law exponent, shape, and scale parameters) could and would be very different for fatigue than for creep rupture and very different for different materials, nevertheless the same mathematical formalism applies in both cases. Necessarily this would be subject to independent experimental verification in any particular case, such as metal fatigue.

A technical area closely related to that of creep rupture and cyclic fatigue is that of size dependence. The disparity between the size of test specimens and the size of full-scale applications could cause divergences—perhaps large divergences. This is a well-known effect, presumably about as difficult to treat as that of life prediction with no size dependence. Some research articles on size/scale dependence are those of Barenblatt [16.20], Barenblatt and Botvina [16.21], and Ritchie [16.22]. The articles by Bazant *et al.* mentioned above give full account of scale dependence, as integrated with life prediction. In the present probabilistic context, size dependence could possibly enter the theory through several of the physical variables, including the shape and scale parameters of the instantaneous static strength as well as the power law form for the kinetic crack growth—especially the exponent in the power law. Many other size-dependent effects could also be involved.

Problem Areas for Study

1. In the creep rupture theory and testing results of Section 16.5, all three systems (carbon–epoxy, glass–epoxy, and aramid–epoxy) have the shape parameters for lifetimes as less than 1, indicating extremely broad distributions, whereas the shape parameters for the static strengths were much larger than 1, indicating a tight grouping of results. The hazard functions for these two drastically different behaviors are very different: the first case is that of a decreasing function with respect to time, while the second is increasing with respect to stress. Are there deeper physical interpretations of these results? Is there some physical requirement that the lifetime shape parameters must be less than 1?
2. Formulate a method and explicit protocol for lifetime testing that minimizes the numbers of required specimens by using the theoretical framework developed in Section 16.5.
3. In terms of strength behavior, are there other physical approaches for deriving the Weibull distribution besides that followed here using the hazard function of reliability theory, or taken from the traditional chain-of-links argument?
4. Relative to a particular type of strength problem, what criteria should be used to decide whether a deterministic formalism is sufficient or if a probabilistic involvement is required?
5. The probabilistic developments here centered on the special significance of the now classical two-parameter Weibull distribution. Does the three-parameter Weibull distribution also have special physical significance, or is it just the usual routine expansion of parameter space.

References

[16.1] Ferry, J. D. (1980) *Viscoelastic Properties of Polymers*, 3rd ed., John Wiley, New York.

[16.2] Christensen, R. M. (2003) *Theory of Viscoelasticity*, 2nd ed., Dover, Mineola, N. Y.

[16.3] Weibull, W. (1951) "A Statistical Distribution Function of Wide Applicability," *J. Applied Mechanics*, **18**, 293–7.

[16.4] Rinne, H. (2009) *The Weibull Distribution*, Chapman & Hall/CRC, Boca Raton, FL.

[16.5] Christensen, R. M., and Miyano, Y. (2006) "Stress Intensity Controlled Kinetic Crack Growth and Stress History Dependent Life Prediction with Statistical Variability," *Int. J. Fracture*, **137**, 77–87.

[16.6] Christensen, R. M., and Miyano, Y. (2007) "Deterministic and Probabilistic Lifetimes from Kinetic Crack Growth: Generalized Forms," *Int. J. Fracture*, **143**, 35–9.

[16.7] Christensen, R. M. and Glaser, R. E. (1994) "Life Prediction for Graphite Epoxy Fiber Composites," *Mechanics Research Comm.*, **21**, 121–30.

[16.8] Christensen, R. M., and Glaser, R. E. (1985) "The Application of Kinetic Fracture Mechanics to Life Prediction for Polymeric Materials," *J. Applied Mechanics*, **52**, 1–5.

[16.9] Glaser, R. E., Christensen, R. M., and Chiao, T. T. (1984) "Theorectical Relations Between Static Strength and Lifetime Distributions for Composites: An Evaluation," *Composites Technology Review*, **6**, 164–7.

[16.10] Christensen, R. M. (2004) "A Probabilistic Treatment of Creep Rupture Behavior for Polymers and Other Materials," *Mechanics Time-Dependent Mats.* **8**, 1–5.

[16.11] Bazant, Z. P., and Pang, S. D. (2007) "Activation Energy Based Extreme Value Statistics and Size Effect in Brittle and Quasibrittle Fracture," *J. Mechanics Physics Solids*, **55**, 91–131.

[16.12] Le, J. L., Bazant, Z. P., and Bazant, M. Z. (2011) "Unified Nano-Mechanics Based Probabilistic Theory of Quasibrittle and Brittle Structures: I. Strength, Static Crack Growth, Lifetime and Scaling," *J. Mechanics Physics Solids*, **59**, 1291–321.

[16.13] Le, J. L., and Bazant, Z. P. (2011) "Unified Nano-Mechanics Based Probabilistic Theory of Quasibrittle and Brittle Structures: II. Fatigue Crack Growth, Lifetime and Scaling," *J. Mechanics Physics Solids*, **59**, 1322–37.

[16.14] Mahesh, S., Phoenix, S. L., and Beyerlein, I. J. (2002) "Strength Distributions and Size Effects for 2D and 3D Composites with Weibull Fibers in an Elastic Matrix," *Int. J. Fracture*, **115**, 41–85.

[16.15] Mahesh, S., and Phoenix, S. L. (2004) "Lifetime Distributions for Unidirectional Fibrous Composites Under Creep-Rupture Loading," *Int. J. Fracture*, **127**, 303–60.

[16.16] Phoenix, S. L., and Newman W. I. (2009) "Time-Dependent Fiber Bundles with Local Load Sharing. II. General Weibull Fibers," *Phys Rev. E.*, **80**, 1–14.

[16.17] Miyano, Y., Nakada, M., and Cai, H. (2008) "Formulation of Long-Term Creep and Fatigue Strengths of Polymer Composites Based on Accelerated Testing Methodology," *J. Composite Materials*, **42**, 1897–919.

[16.18] Nakada, M., and Miyano, Y. (2009) "Accelerated Testing for Long-Term Fatigue Strength of Various FRP Laminates for Marine Use," *Composites Sciencce and Technology*, **69**, 805–13.

[16.19] Noda, J., Nakada, M., and Miyano, Y. (2009) "Statistical Scatter of Creep and Fatigue Failure Times for CFRP Laminates," *J. Reinforced Plastics and Composites*, **28**, 1139–48.

[16.20] Barenblatt, G. I. (2003) *Scaling*, Cambridge University Press, Cambridge, UK.

[16.21] Barenblatt, G. I., and Botvina, L. R. (1981) "Incomplete Self-Similarity of Fatigue in the Linear Range of Crack Growth," *Fatigue of Engrg. Mater. & Structures*, **3**, 193–202.

[16.22] Ritchie, R. O. (2005) "Incomplete Self-Similarity and Fatigue-Crack Growth," *Int. J. Fracture*, **132**, 197–203.

This page intentionally left blank

Index